GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:11:58 ; Search time 22.75 Seconds
(without alignments)
2048.317 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2242_2332
Perfect score: 484
Sequence: 1 VNNPKEWLQVDFQKTMKVTG.....VNSLDPPLLTRYLRIHPQSW 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	484	100.0	216	2	Q14286	Q14286 homo sapien
2	484	100.0	2351	1	FA8_HUMAN	P00451 homo sapien
3	417	86.2	2343	2	O18806	O18806 canis famil
4	412	85.1	2343	2	O62730	O62730 canis famil
5	406	83.9	2319	1	FA8_MOUSE	Q06194 mus musculu
6	392	81.0	2133	1	FA8_PIG	P12263 sus scrofa
7	351	72.5	2258	2	Q7TN96	Q7tn96 rattus norv
8	291	60.1	1377	2	Q804X3	Q804x3 gallus gall
9	240	49.6	1639	2	Q804W6	Q804w6 fugu rubrip
10	208	43.0	427	1	MFGM_RAT	P70490 rattus norv
11	206	42.6	116	2	Q8C7H9	Q8c7h9 mus musculu
12	206	42.6	2183	2	O88783	O88783 mus musculu
13	205	42.4	463	1	MFGM_MOUSE	P21956 mus musculu
14	204	42.1	2102	2	Q7TPK2	Q7tpk2 rattus norv
15	203	41.9	745	2	Q804X4	Q804x4 gallus gall
16	202	41.7	2258	1	FA5_PIG	Q9glp1 sus scrofa
17	199	41.1	2224	1	FA5_HUMAN	P12259 homo sapien
18	199	41.1	2224	2	Q6UPU6	Q6upu6 homo sapien
19	198	40.9	481	2	Q6RUW2	Q6ruw2 xenopus lae
20	198	40.9	2211	1	FA5_BOVIN	Q28107 bos taurus
21	197	40.7	384	2	Q8C8K0	Q8c8k0 mus musculu
22	197	40.7	470	2	Q8C4U8	Q8c4u8 mus musculu
23	197	40.7	480	1	EDI3_MOUSE	O35474 mus musculu
24	197	40.7	480	2	Q8CBF7	Q8cbf7 mus musculu
25	196	40.5	480	1	EDI3_HUMAN	O43854 homo sapien
26	196	40.5	480	2	Q8N610	Q8n610 homo sapien
27	196	40.5	1460	2	Q7SZN0	Q7szn0 pseudonaja
28	194	40.1	407	2	Q8AYE0	Q8aye0 brachydanio
29	194	40.1	2119	2	Q90X47	Q90x47 brachydanio
30	191.5	39.6	858	2	O76470	O76470 lytechinus
31	190	39.3	475	2	Q68EW5	Q68ew5 xenopus lae
32	190	39.3	764	1	CPX2_MOUSE	Q9d215 mus musculu
33	189	39.0	475	2	Q6P1V9	Q6p1v9 xenopus tro
34	189	39.0	590	2	Q658N0	Q658n0 homo sapien
35	189	39.0	756	1	CPX2_HUMAN	Q8n436 homo sapien
36	189	39.0	769	1	DCB2_MOUSE	Q91zv3 mus musculu
37	187	38.6	715	1	DCB1_HUMAN	Q8n8z6 homo sapien
38	186	38.4	769	1	DCB2_RAT	Q91zv2 rattus norv
39	186	38.4	775	1	DCB2_HUMAN	Q96pd2 homo sapien
40	181	37.4	1802	2	Q804W5	Q804w5 fugu rubrip
41	180	37.2	925	1	NRP2_RAT	O35276 rattus norv
42	178	36.8	728	2	O97567	O97567 bos taurus
43	177	36.6	555	2	Q9H2E2	Q9h2e2 homo sapien
44	177	36.6	901	2	Q9H2D5	Q9h2d5 homo sapien
45	177	36.6	901	2	Q9H2E4	Q9h2e4 homo sapien

ALIGNMENTS

RESULT 1
Q14286
ID Q14286 PRELIMINARY; PRT; 216 AA.
AC Q14286;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Coagulation factor VIII associated protein B (Coagulation factor VIII,
DE isoform b).
GN Name=F8B;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93052386; PubMed=1427887;
RA Levinson B., Kenwrick S., Gamel P., Fisher K., Gitschier J.;
RT "Evidence for a third transcript from the human factor VIII gene.";
RL Genomics 14:585-589(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.

```
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; M90707; AAA58466.1; -.
DR   EMBL; BC022513; AAH22513.1; -.
DR   EMBL; BC064380; AAH64380.1; -.
DR   PIR; A44258; A44258.
DR   HSSP; P00451; 1CFG.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 1.
DR   PROSITE; PS01285; FA58C_1; 1.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
SQ   SEQUENCE   216 AA;  24641 MW;  6C82D4F89E35A376 CRC64;

  Query Match           100.0%;  Score 484;  DB 2;  Length 216;
  Best Local Similarity  100.0%;  Pred. No. 5e-45;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 166

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db        167 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 197


RESULT 2
FA8_HUMAN
ID   FA8_HUMAN      STANDARD;      PRT;  2351 AA.
AC   P00451;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Coagulation factor VIII precursor (Procoagulant component)
DE   (Antihemophilic factor) (AHF).
GN   Name=F8; Synonyms=F8C;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86081164; PubMed=3935400;
RA   Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,
RA   Hartog K., Kuo C.H., Masiarz F.R., Merryweather J.P., Najarian R.,
RA   Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,
RA   Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,
RA   Nordfang O., Ezban M.;
RT   "Characterization of the polypeptide composition of human factor
RT   VIII:C and the nucleotide sequence and expression of the human kidney
RT   cDNA.";
RL   DNA 4:333-349(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85061548; PubMed=6438526;
RA   Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,
RA   Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,
RA   Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;
RT   "Expression of active human factor VIII from recombinant DNA clones.";
RL   Nature 312:330-337(1984).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85061550; PubMed=6438528;
RA   Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,
RA   Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,
RA   Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,
RA   Hewick R.M.;
RT   "Molecular cloning of a cDNA encoding human antihaemophilic factor.";
RL   Nature 312:342-347(1984).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93265012; PubMed=1303178;
RA   Gitschier J., Wood W.I.;
RT   "Sequence of the exon-containing regions of the human factor VIII
RT   gene.";
RL   Hum. Mol. Genet. 1:199-200(1992).
RN   [5]
RP   SEQUENCE OF 2064-2070 FROM N.A.
RA   de Water N.S., Williams R., Browett P.J.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SULFATION OF TYR-1699.
RX   MEDLINE=91093266; PubMed=1898735;
RA   Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
RA   Mertens K., van Mourik J.A.;
RT   "Sulfation of Tyr1680 of human blood coagulation factor VIII is
RT   essential for the interaction of factor VIII with von Willebrand
RT   factor.";
RL   J. Biol. Chem. 266:740-746(1991).
RN   [7]
RP   SULFATION.
RX   MEDLINE=92207952; PubMed=1554716;
RA   Pittman D.D., Wang J.H., Kaufman R.J.;
RT   "Identification and functional importance of tyrosine sulfate residues
RT   within recombinant factor VIII.";
RL   Biochemistry 31:3315-3325(1992).
RN   [8]
RP   SULFATION SITES TYR-737; TYR-738 AND TYR 742, AND DISULFIDE BONDS.
RX   PubMed=7613471;
RA   McMullen B.A., Fujikawa K., Davie E.W., Hedner U., Ezban M.;
RT   "Locations of disulfide bonds and free cysteines in the heavy and
RT   light chains of recombinant human factor VIII (antihemophilic factor
RT   A).";
RL   Protein Sci. 4:740-746(1995).
RN   [9]
RP   STRUCTURE BY NMR OF 2322-2343.
RX   MEDLINE=95200924; PubMed=7893714;
RA   Gilbert G.E., Baleja J.D.;
RT   "Membrane-binding peptide from the C2 domain of factor VIII forms an
RT   amphipathic structure as determined by NMR spectroscopy.";
RL   Biochemistry 34:3022-3031(1995).
RN   [10]
RP   REVIEW ON MOLECULAR BASIS OF HEMA.
RX   MEDLINE=91221499; PubMed=1902642;
RA   Gitschier J.;
RT   "The molecular basis of hemophilia A.";
RL   Ann. N. Y. Acad. Sci. 614:89-96(1991).
RN   [11]
RP   REVIEW ON MOLECULAR BASIS OF HEMA.
RX   MEDLINE=89088506; PubMed=2491949;
RA   White G.C. II, Shoemaker C.B.;
RT   "Factor VIII gene and hemophilia A.";
RL   Blood 73:1-12(1989).
RN   [12]
RP   REVIEW ON MOLECULAR BASIS OF HEMA.
RX   MEDLINE=95245332; PubMed=7728145;
RA   Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
RT   "Molecular etiology of factor VIII deficiency in hemophilia A.";
RL   Hum. Mutat. 5:1-22(1995).
RN   [13]
RP   VARIANT HEMA GLN-2326.
RX   MEDLINE=86235434; PubMed=3012775;
RA   Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
RT   "Identification of a missense mutation in the factor VIII gene of a
RT   mild hemophiliac.";
RL   Science 232:1415-1416(1986).
RN   [14]
RP   VARIANT HEMA PRO-2135.
RX   MEDLINE=88096539; PubMed=3122181;
RA   Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
RT   "A novel missense mutation in the factor VIII gene identified by
RT   analysis of amplified hemophilia DNA sequences.";
RL   Nucleic Acids Res. 15:9797-9805(1987).
```

```
RN   [15]
RP   VARIANT HEMA GLN-2228.
RX   MEDLINE=88191889; PubMed=2833855;
RA   Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
RA   Kazazian H.H.;
RT   "Nonsense and missense mutations in hemophilia A: estimate of the
RT   relative mutation rate at CG dinucleotides.";
RL   Am. J. Hum. Genet. 42:718-725(1988).
RN   [16]
RP   VARIANT HEMA GLY-291.
RX   MEDLINE=88220354; PubMed=2835904;
RA   Youssoufian H., Wong C., Aronis S., Platokoukis H., Kazazian H.H. Jr.,
RA   Antonarakis S.E.;
RT   "Moderately severe hemophilia A resulting from Glu-->Gly substitution
RT   in exon 7 of the factor VIII gene.";
RL   Am. J. Hum. Genet. 42:867-871(1988).
RN   [17]
RP   VARIANT HEMA CYS-1708.
RX   MEDLINE=89274393; PubMed=2499363;
RA   O'Brien D.P., Tuddenham E.G.;
RT   "Purification and characterization of factor VIII 1,689-Cys: a
RT   nonfunctional cofactor occurring in a patient with severe hemophilia
RT   A.";
RL   Blood 73:2117-2122(1989).
RN   [18]
RP   VARIANT HEMA CYS-391.
RX   MEDLINE=90001543; PubMed=2506948;
RA   Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
RT   "An arginine to cysteine amino acid substitution at a critical
RT   thrombin cleavage site in a dysfunctional factor VIII molecule.";
RL   Blood 74:1612-1617(1989).
RN   [19]
RP   VARIANT HEMA LEU-189.
RX   MEDLINE=90057680; PubMed=2510835;
RA   Chan V., Chan T.K., Tong T.M., Todd D.;
RT   "A novel missense mutation in exon 4 of the factor VIII:C gene
RT   resulting in moderately severe hemophilia A.";
RL   Blood 74:2688-2691(1989).
RN   [20]
RP   VARIANT HEMA LEU-2326.
RX   MEDLINE=89197216; PubMed=2495245;
RA   Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
RT   "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
RT   of the factor VIII gene.";
RL   Hum. Genet. 81:335-338(1989).
RN   [21]
RP   VARIANT HEMA HIS-391.
RX   MEDLINE=89264602; PubMed=2498882;
RA   Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
RA   Fujimaki M., Hoyer L.W.;
RT   "Direct characterization of factor VIII in plasma: detection of a
RT   mutation altering a thrombin cleavage site (arginine-
RT   372-->histidine).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
RN   [22]
RP   VARIANT HEMA CYS-1708.
RX   MEDLINE=90105723; PubMed=2104766;
RA   Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
RA   Phillips J.A. III, Janco R.L., Hoyer L.W.;
RT   "Characterization of a thrombin cleavage site mutation (Arg 1689 to
RT   Cys) in the factor VIII gene of two unrelated patients with cross-
RT   reacting material-positive hemophilia A.";
RL   Blood 75:384-389(1990).
RN   [23]
RP   VARIANTS HEMA GLN-2228 AND LEU-2326.
RX   MEDLINE=90123183; PubMed=2105106;
RA   Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
RA   Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
RT   "Recurrent mutations and three novel rearrangements in the factor VIII
RT   gene of hemophilia A patients of Italian descent.";
RL   Blood 75:662-670(1990).
RN   [24]
RP   VARIANT HEMA CYS-391.
RX   MEDLINE=90329422; PubMed=1973901;
RA   Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
RT   "CRM+ haemophilia A due to a missense mutation (372-->Cys) at the
RT   internal heavy chain thrombin cleavage site.";
RL   Br. J. Haematol. 75:73-77(1990).
RN   [25]
RP   VARIANTS HEMA PHE-1699 AND CYS-1708.
RX   MEDLINE=90152691; PubMed=2105906;
RA   Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
RA   Kazazian H.H., Antonarakis S.E.;
RT   "Characterization of mutations in the factor VIII gene by direct
RT   sequencing of amplified genomic DNA.";

  Query Match             100.0%;  Score 484;  DB 1;  Length 2351;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-44;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2242 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2301

Qy     61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
          |||||||||||||||||||||||||||||||
Db   2302 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2332


RESULT 3
O18806
ID   O18806      PRELIMINARY;      PRT;  2343 AA.
AC   O18806;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Factor VIII.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Cameron C., Notley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
RA   Giles A., Lillicrap D.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF016234; AAB87412.1; -.
DR   HSSP; P00451; 1D7P.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ   SEQUENCE   2343 AA;  265829 MW;  A854FAE571C3B399 CRC64;

  Query Match             86.2%;  Score 417;  DB 2;  Length 2343;
  Best Local Similarity   84.4%;  Pred. No. 1.7e-36;
  Matches   76;  Conservative    7;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVF 61
          |||||||||||:|||||:|||||||| |||||||||||||| |||| ||||||||
Db    2235 NNPKEWLQVDFRKTMKVTGITTQGVKSLLISMYVKEFLISSSQDGHNWTLFLQNGKVKVF 2294

Qy      62 QGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
           |||:|| ||| | |:|||: ||:|:|||||
Db    2295 QGNRDSSTPVRNRLEPPLVARYVRLHPQSW 2324
```

```
RESULT 4
O62730
ID   O62730      PRELIMINARY;      PRT;  2343 AA.
AC   O62730;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Factor VIII.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Gordy P.W., Bowen R.A.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF049489; AAC05384.1; -.
DR   HSSP; P00451; 1IQD.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ   SEQUENCE   2343 AA;  265613 MW;  F612D744ADAADD99 CRC64;

  Query Match             85.1%;  Score 412;  DB 2;  Length 2343;
  Best Local Similarity   83.3%;  Pred. No. 6.2e-36;
  Matches   75;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVF 61
              |||||||||||:|||||||:|||||||| |||||||||||||||| |||| || |||||
Db       2235 NNPKEWLQVDFRKTMKVTGITTQGVKSLLISMYVKEFLISSSQDGHNWTLFLQNDKVKVF 2294

Qy         62 QGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||:|| ||| |:|:|||: ||:|:|||||
Db       2295 QGNRDSSTPVRNALEPPLVARYVRLHPQSW 2324


RESULT 5
FA8_MOUSE
ID   FA8_MOUSE      STANDARD;      PRT;  2319 AA.
AC   Q06194;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Coagulation factor VIII precursor (Procoagulant component).
GN   Name=F8; Synonyms=Cf8, F8c;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX   MEDLINE=93300511; PubMed=8314577;
RA   Elder B., Lakich D., Gitschier J.;
RT   "Sequence of the murine factor VIII cDNA.";
RL   Genomics 16:374-379(1993).
CC   -!- FUNCTION: Factor VIII, along with calcium and phospholipid, acts
CC       as a cofactor for factor IXa when it converts factor X to the
CC       activated form, factor Xa.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Found in most tissues.
CC   -!- PTM: The binding of Von Willibrand Factor (VWF) and activation
CC       depend on the sulfation of Tyr-1669.
CC   -!- SIMILARITY: Belongs to the multicopper oxidase family.
CC   -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L05573; AAA37385.1; -.
DR   PIR; A47004; A47004.
DR   HSSP; P00451; 1D7P.
DR   MGD; MGI:88383; F8.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW   Acute phase; Blood coagulation; Calcium; Glycoprotein; Plasma; Repeat;
KW   Signal; Sulfation.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20   2319       Coagulation factor VIII.
FT   DOMAIN       20    349       F5/8 type A 1.
FT   DOMAIN       20    199       Plastocyanin-like 1.
FT   DOMAIN      207    349       Plastocyanin-like 2.
FT   DOMAIN      399    730       F5/8 type A 2.
FT   DOMAIN      399    573       Plastocyanin-like 3.
FT   DOMAIN      583    730       Plastocyanin-like 4.
FT   DOMAIN      760   1640       B.
FT   DOMAIN     1683   2008       F5/8 type A 3.
FT   DOMAIN     1683   1845       Plastocyanin-like 5.
FT   DOMAIN     1855   2008       Plastocyanin-like 6.
FT   DOMAIN     2008   2156       F5/8 type C 1.
FT   DOMAIN     2161   2313       F5/8 type C 2.
FT   SITE        391    392       Cleavage (by thrombin) (By similarity).
FT   SITE        759    760       Cleavage (by thrombin) (By similarity).
FT   SITE       1324   1325       Cleavage (activation) (By similarity).
FT   SITE       1640   1641       Cleavage (activation) (By similarity).
FT   SITE       1678   1679       Cleavage (by thrombin) (By similarity).
FT   MOD_RES     367    367       Sulfotyrosine (By similarity).
FT   MOD_RES     737    737       Sulfotyrosine (By similarity).
FT   MOD_RES     738    738       Sulfotyrosine (By similarity).
FT   MOD_RES     742    742       Sulfotyrosine (By similarity).
FT   MOD_RES    1669   1669       Sulfotyrosine (By similarity).
FT   MOD_RES    1687   1687       Sulfotyrosine (By similarity).
FT   DISULFID    173    199       Probable.
FT   DISULFID    547    573       Probable.
FT   DISULFID   1819   1845       Probable.
FT   DISULFID   2008   2156       By similarity.
FT   DISULFID   2161   2313       By similarity.
FT   CARBOHYD     61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    259    259       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    423    423       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    601    601       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    880    880       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    958    958       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1015   1015       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1022   1022       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1026   1026       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1044   1044       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1076   1076       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1087   1087       N-linked (GlcNAc. . .) (Potential).
```

```
FT   CARBOHYD   1136   1136       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1161   1161       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1192   1192       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1255   1255       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1268   1268       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1273   1273       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1274   1274       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1302   1302       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1316   1316       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1340   1340       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1378   1378       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1797   1797       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   2105   2105       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   2319 AA;  266148 MW;  FD054DE051DB2A01 CRC64;

  Query Match             83.9%;  Score 406;  DB 1;  Length 2319;
  Best Local Similarity   82.4%;  Pred. No. 2.8e-35;
  Matches   75;  Conservative    6;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||:||:||||| |||||||: ||||||| |||:||||||||||||| ||    ||||||
Db       2210 VNDPKQWLQVDLQKTMKVTGIITQGVKSLFTSMFVKEFLISSSQDGHHWTQILYNGKVKV 2269

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||| ||::|||||||||||||||| |
Db       2270 FQGNQDSSTPMMNSLDPPLLTRYLRIHPQIW 2300


RESULT 6
FA8_PIG
ID   FA8_PIG        STANDARD;      PRT;  2133 AA.
AC   P12263; Q95243;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Coagulation factor VIII precursor (Procoagulant component).
GN   Name=F8; Synonyms=CF8;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Healey J.F., Lubin I.M., Lollar P.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 705-1573 FROM N.A.
RX   MEDLINE=86287369; PubMed=3016730;
RA   Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
RA   Kaufman R.J.;
RT   "A large region (approximately equal to 95 kDa) of human factor VIII
RT   is dispensable for in vitro procoagulant activity.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
RN   [3]
RP   SEQUENCE OF 392-759 FROM N.A.
RX   MEDLINE=94179260; PubMed=7510693;
RA   Lubin I.M., Healey J.F., Scandella D., Runge M.S., Lollar P.;
RT   "Elimination of a major inhibitor epitope in factor VIII.";
RL   J. Biol. Chem. 269:8639-8641(1994).
CC   -!- FUNCTION: Factor VIII, along with calcium and phospholipid, acts
CC       as a cofactor for factor IXa when it converts factor X to the
CC       activated form, factor Xa.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- SIMILARITY: Belongs to the multicopper oxidase family.
CC   -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U49517; AAB06705.1; -.
DR   PIR; A25945; A25945.
DR   PIR; T42763; T42763.
DR   HSSP; P00451; 1D7P.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW   Acute phase; Blood coagulation; Calcium; Glycoprotein; Plasma; Repeat;
KW   Signal; Sulfation.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20   2133       Coagulation factor VIII.
FT   DOMAIN       20    357       F5/8 type A 1.
FT   DOMAIN       20    199       Plastocyanin-like 1.
FT   DOMAIN      207    357       Plastocyanin-like 2.
FT   DOMAIN      399    730       F5/8 type A 2.
FT   DOMAIN      399    573       Plastocyanin-like 3.
FT   DOMAIN      583    730       Plastocyanin-like 4.
FT   DOMAIN      760   1599       B.
FT   DOMAIN     1495   1822       F5/8 type A 3.
FT   DOMAIN     1495   1659       Plastocyanin-like 5.
FT   DOMAIN     1669   1822       Plastocyanin-like 6.
FT   DOMAIN     1822   1970       F5/8 type C 1.
FT   DOMAIN     1975   2127       F5/8 type C 2.
FT   SITE        391    392       Cleavage (by thrombin) (By similarity).
FT   SITE        759    760       Cleavage (by thrombin) (By similarity).
FT   SITE       1449   1450       Cleavage (activation) (By similarity).
FT   SITE       1490   1491       Cleavage (by thrombin) (By similarity).
FT   MOD_RES     737    737       Sulfotyrosine (By similarity).
FT   MOD_RES     738    738       Sulfotyrosine (By similarity).
FT   MOD_RES     742    742       Sulfotyrosine (By similarity).
FT   DISULFID    173    199       Probable.
FT   DISULFID    547    573       Probable.
FT   DISULFID   1633   1659       Probable.
FT   DISULFID   1822   1970       By similarity.
FT   DISULFID   1975   2127       By similarity.
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    259    259       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    601    601       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    929    929       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    985    985       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1025   1025       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1111   1111       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1181   1181       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1208   1208       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1245   1245       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1265   1265       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1335   1335       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1408   1408       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1611   1611       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1919   1919       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT    713    713       N -> M (in Ref. 2).
FT   CONFLICT    734    734       I -> T (in Ref. 2).
FT   CONFLICT    792    792       G -> Q (in Ref. 2).
FT   CONFLICT   1133   1133       E -> F (in Ref. 2).
FT   CONFLICT   1191   1191       I -> L (in Ref. 2).
FT   CONFLICT   1209   1209       R -> F (in Ref. 2).
FT   CONFLICT   1437   1437       C -> G (in Ref. 2).
FT   CONFLICT   1456   1456       F -> R (in Ref. 2).
FT   CONFLICT   1539   1539       F -> R (in Ref. 2).
FT   CONFLICT   1546   1546       Q -> N (in Ref. 2).
SQ   SEQUENCE   2133 AA;  239304 MW;  152BBA8997F570DA CRC64;

  Query Match             81.0%;  Score 392;  DB 1;  Length 2133;
```

```
  Best Local Similarity   79.1%;  Pred. No. 8.8e-34;
  Matches   72;  Conservative   10;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              |:: :|||||| |||:||||:||||||||||:||||||||:|||||| :|||| |:|  ||
Db       2024 VSSAEEWLQVDLQKTVKVTGITTQGVKSLLSSMYVKEFLVSSSQDGRRWTLFLQDGHTKV 2083

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||| |||||:||||| |||||||| ||
Db       2084 FQGNQDSSTPVVNALDPPLFTRYLRIHPTSW 2114


RESULT 7
Q7TN96
ID   Q7TN96      PRELIMINARY;      PRT;  2258 AA.
AC   Q7TN96;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Factor VIII.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RA   Watzka M., Geisen C., Seifried E., Oldenburg J.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AY362193; AAQ21580.1; -.
DR   HSSP; P00451; 1CFG.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ   SEQUENCE   2258 AA;  251299 MW;  13AF91C788059B1D CRC64;

  Query Match             72.5%;  Score 351;  DB 2;  Length 2258;
  Best Local Similarity   69.2%;  Pred. No. 3.1e-29;
  Matches   63;  Conservative   14;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||:| :||||| |:|:|||||| ||| :||||:|:||:||:|:|||| ||   |:|||||
Db       2149 VNDPTQWLQVDLQRTVKVTGVVTQGARSLLTAMFVKKFLVSTSQDGRHWTHVLQDGKVKV 2208

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||:|: ||:|||| ||  |||||||| |
Db       2209 FQGNRDASTPMVNSLHPPRFTRYLRIHPQVW 2239


RESULT 8
Q804X3
ID   Q804X3      PRELIMINARY;      PRT;  1377 AA.
AC   Q804X3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Coagulation factor VIII (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA   Tuddenham E.G.D., McVey J.H.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF465272; AAO33367.1; -.
DR   HSSP; P00451; 1D7P.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 1.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   1377 AA;  152728 MW;  15ACFD7F765665CD CRC64;

  Query Match             60.1%;  Score 291;  DB 2;  Length 1377;
  Best Local Similarity   55.6%;  Pred. No. 7.1e-23;
  Matches   50;  Conservative   17;  Mismatches   23;  Indels    0;  Gaps    0;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVF 61
              |:| |||||||: | ||| : ||| |:: | |:|||| :||:||| |:    ||| |:|
Db       1269 NSPSEWLQVDFEATKKVTAIITQGAKAVFTHMFVKEFAVSSNQDGVHWSRVLHNGKEKIF 1328

Qy         62 QGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              : |:|  : |:|||:|||  ||:||||: |
Db       1329 RANRDYTSTVLNSLEPPLFARYVRIHPRHW 1358


RESULT 9
Q804W6
ID   Q804W6      PRELIMINARY;      PRT;  1639 AA.
AC   Q804W6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Coagulation factor VIII.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA   Tuddenham E.G.D., McVey J.H.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF465279; AAO33374.1; -.
DR   HSSP; P00451; 1CFG.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_2.
SQ   SEQUENCE   1639 AA;  185203 MW;  881778AAFF809006 CRC64;

  Query Match             49.6%;  Score 240;  DB 2;  Length 1639;
  Best Local Similarity   48.9%;  Pred. No. 3.5e-17;
  Matches   45;  Conservative   12;  Mismatches   33;  Indels    2;  Gaps    1;
```

```
Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNG--KVK 59
              ||| ||||||  |  ::||| ||| :|||| | | || :: |:||  |:   :    : |
Db       1529 NNPHEWLQVDLGKVKRITGVVTQGARSLLTKMMVTEFSVTISRDGQAWSSVLEGSSQREK 1588

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :|||| ||    :   | ||  ||:|||| |
Db       1589 IFQGNNDSDEEALTIFDAPLFGRYIRIHPLGW 1620


RESULT 10
MFGM_RAT
ID   MFGM_RAT       STANDARD;      PRT;   427 AA.
AC   P70490;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Lactadherin precursor (Milk fat globule-EGF factor 8) (MFG-E8) (O-
DE   acetyl GD3 ganglioside synthase) (AGS) (MFGM).
GN   Name=Mfge8; Synonyms=Ags;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=96374422; PubMed=8780713; DOI=10.1006/bbrc.1996.1274;
RA   Ogura K., Nara K., Watanabe Y., Kohno K., Tai T., Sanai Y.;
RT   "Cloning and expression of cDNA for O-acetylation of GD3
RT   ganglioside.";
RL   Biochem. Biophys. Res. Commun. 225:932-938(1996).
CC   -!- FUNCTION: Specific ligand for the alpha-v/beta-3 and alpha-v/beta-
CC       5 receptors. Also binds to phosphatidylserine-enriched cell
CC       surfaces in a receptor-independent manner. Zona pellucida-binding
CC       protein which may play a role in gamete interaction. Appears to
CC       participate in the O-acetylation of GD3 ganglioside sialic acid.
CC   -!- SUBCELLULAR LOCATION: Peripheral membrane protein.
CC   -!- TISSUE SPECIFICITY: Spleen, lung, heart, brain and muscle.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D84068; BAA12210.1; -.
DR   PIR; JC4915; JC4915.
DR   HSSP; P00740; 1EDM.
DR   RGD; 3083; Mfge8.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
KW   Cell adhesion; EGF-like domain; Fertilization; Glycoprotein; Repeat;
KW   Signal.
FT   SIGNAL        1     22       Potential.
FT   CHAIN        23    427       Lactadherin.
FT   DOMAIN       24     61       EGF-like 1.
FT   DOMAIN       64    108       EGF-like 2.
FT   DOMAIN      111    267       F5/8 type C 1.
FT   DOMAIN      272    427       F5/8 type C 2.
FT   DISULFID     28     39       By similarity.
FT   DISULFID     33     49       By similarity.
FT   DISULFID     51     60       By similarity.
FT   DISULFID     68     79       By similarity.
FT   DISULFID     73     96       By similarity.
FT   DISULFID     98    107       By similarity.
FT   DISULFID    111    267       By similarity.
FT   DISULFID    254    258       By similarity.
FT   DISULFID    272    427       By similarity.
FT   SITE         87     89       Cell attachment site (Potential).
FT   CARBOHYD     61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    230    230       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    280    280       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    390    390       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   427 AA;  47413 MW;  EA8C8631F3EE6047 CRC64;

  Query Match             43.0%;  Score 208;  DB 1;  Length 427;
  Best Local Similarity   43.3%;  Pred. No. 2.6e-14;
  Matches   39;  Conservative   15;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVF 61
              |: |||||||     ||||: ||| :          ||   : :: | || |||:: : |  |||
Db        325 NSAKEWLQVDLGTQKKVTGIITQGARDFGHIQYVASYKVAHSDDGVQWTVYEEQGTSKVF 384

Qy         62 QGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||| |: :   |  : | : ||:|: | ||
Db        385 QGNLDNNSHKKNIFEKPFMARYVRVLPLSW 414


RESULT 11
Q8C7H9
ID   Q8C7H9      PRELIMINARY;      PRT;   116 AA.
AC   Q8C7H9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mus musculus adult male liver tumor cDNA, RIKEN full-length enriched
DE   library, clone:C730026K03 product:coagulation factor V, full insert
DE   sequence. (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
```

```
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Liver;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 F5/8 type C domain.
DR   EMBL; AK050205; BAC34121.1; -.
DR   HSSP; P12259; 1CZT.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00754; F5_F8_type_C; 1.
DR   SMART; SM00231; FA58C; 1.
DR   PROSITE; PS01285; FA58C_1; 1.
DR   PROSITE; PS01286; FA58C_2; 1.
DR   PROSITE; PS50022; FA58C_3; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   116 AA;  13436 MW;  FDCDF8E0C65EB1F1 CRC64;

  Query Match             42.6%;  Score 206;  DB 2;  Length 116;
  Best Local Similarity   42.4%;  Pred. No. 9.8e-15;
  Matches   39;  Conservative   19;  Mismatches   32;  Indels    2;  Gaps    1;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--K 59
              || |:|||||  |  ||| : ||| ||| : |||| : |  |  |  |  : |   :  |
Db          9 NNNKQWLQVDLLKIKKVTAIVTQGCKSLSSEMYVKSYSIQYSDQGVAWKPYRQKSSMVDK 68

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :|:|| ::   : |  :||:::|::|| |::|
Db         69 IFEGNSNTKGHMKNFFNPPIISRFIRIIPKTW 100


RESULT 12
O88783
ID   O88783      PRELIMINARY;      PRT;  2183 AA.
AC   O88783;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Coagulation factor V.
GN   Name=F5;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98282202; PubMed=9616155;
RA   Yang T.L., Cui J., Rehumtulla A., Yang A., Moussalli M., Kaufman R.J.,
RA   Ginsburg D.;
RT   "The structure and function of murine factor V and its inactivation by
RT   protein C.";
RL   Blood 91:4593-4599(1998).
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; U52925; AAC99553.1; -.
DR   PIR; T42764; T42764.
DR   HSSP; P12259; 1CZT.
DR   MGD; MGI:88382; F5.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR009271; LSPD.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   Pfam; PF06049; LSPR; 24.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_2.
SQ   SEQUENCE   2183 AA;  247228 MW;  BF0A8AA723F60317 CRC64;

  Query Match             42.6%;  Score 206;  DB 2;  Length 2183;
  Best Local Similarity   42.4%;  Pred. No. 2.7e-13;
  Matches   39;  Conservative   19;  Mismatches   32;  Indels    2;  Gaps    1;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--K 59
              || |:|||||  |  ||| : ||| ||| : |||| : |  |  |  |  : |   :  |
Db       2076 NNNKQWLQVDLLKIKKVTAIVTQGCKSLSSEMYVKSYSIQYSDQGVAWKPYRQKSSMVDK 2135

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :|:|| ::   : |  :||:::|::|| |::|
Db       2136 IFEGNSNTKGHMKNFFNPPIISRFIRIIPKTW 2167


RESULT 13
MFGM_MOUSE
ID   MFGM_MOUSE     STANDARD;      PRT;   463 AA.
AC   P21956; P97800; Q9R1X9; Q9WTS3;
DT   01-AUG-1991 (Rel. 19, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Lactadherin precursor (Milk fat globule-EGF factor 8) (MFG-E8) (MFGM)
DE   (Sperm surface protein SP47) (MP47).
GN   Name=Mfge8;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), SEQUENCE OF 23-35, TISSUE SPECIFICITY,
RP   SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RC   STRAIN=BALB/c; TISSUE=Mammary gland;
RX   MEDLINE=91046008; PubMed=2122462;
RA   Stubbs J.D., Lekutis C., Singer K.L., Bui A., Yuzuki D.,
RA   Srinivasan U., Parry G.;
RT   "cDNA cloning of a mouse mammary epithelial cell surface protein
RT   reveals the existence of epidermal growth factor-like domains linked
RT   to factor VIII-like sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:8417-8421(1990).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RA   Ensslin M.A.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 23-463 FROM N.A. (ISOFORM 2).
RC   TISSUE=Testis;
RX   MEDLINE=98206817; PubMed=9546740;
RA   Ensslin M.A., Vogel T., Calvete J.J., Thole H.H., Schmidtke J.,
RA   Matsuda T., Toepfer-Petersen E.;
```

```
RT   "Molecular cloning and characterization of P47, a novel boar sperm-
RT   associated zona pellucida-binding protein homologous to a family of
RT   mammalian secretory proteins.";
RL   Biol. Reprod. 58:1057-1064(1998).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), TISSUE SPECIFICITY,
RP   DEVELOPMENTAL STAGE, INDUCTION, AND GLYCOSYLATION.
RC   STRAIN=BALB/c; TISSUE=Mammary gland;
RX   MEDLINE=99120894; PubMed=9920772; DOI=10.1006/bbrc.1998.0107;
RA   Oshima K., Aoki N., Negi M., Kishi M., Kitajima K., Matsuda T.;
RT   "Lactation-dependent expression of an mRNA splice variant with an exon
RT   for a multiply O-glycosylated domain of mouse milk fat globule
RT   glycoprotein MFG-E8.";
RL   Biochem. Biophys. Res. Commun. 254:522-528(1999).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=NOD; TISSUE=Dendritic cell;
RX   MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA   Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA   Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA   Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA   Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA   Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA   Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA   Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA   Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA   Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA   Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA   Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA   Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA   Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA   Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA   Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA   Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA   Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA   Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA   Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA   Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA   Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA   Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA   Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA   Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA   Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA   Birney E., Hayashizaki Y.;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Specific ligand for the alpha-v/beta-3 and alpha-v/beta-
CC       5 receptors. Also binds to phosphatidylserine-enriched cell
CC       surfaces in a receptor-independent manner. Zona pellucida-binding
CC       protein which may play a role in gamete interaction (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Peripheral membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1; Synonyms=Long;
CC         IsoId=P21956-1; Sequence=Displayed;
CC       Name=2; Synonyms=Short;
CC         IsoId=P21956-2; Sequence=VSP_009880;
CC   -!- TISSUE SPECIFICITY: Mammary epithelial cell surfaces and
CC       spermatozoan. Isoform 2 is present in brain, heart, kidney and
CC       spleen and at low levels in lung, liver, small intestine and
CC       testis.
CC   -!- DEVELOPMENTAL STAGE: Isoform 1 and isoform 2 are detectable in
CC       mammary tissue from non-pregnant animals, with isoform 2 being
CC       predominant. Levels of isoform 1 increase during gestation and
CC       lactation while levels of isoform 2 decrease.
CC   -!- INDUCTION: Isoform 1 is induced by insulin, prolactin and
CC       hydrocortisone in mammary epithelial cells. Expression of isoform
CC       2 is repressed by the same treatment.
CC   -!- PTM: N-glycosylated. Isoform 1 also exists in both an O-
CC       glycosylated and a non-O-glycosylated form.
CC   -!- SIMILARITY: Contains 2 EGF-like domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M38337; AAA39534.1; -.
DR   EMBL; Y11684; CAA72380.2; -.
DR   EMBL; AB021130; BAA35180.1; -.
DR   EMBL; AB025280; BAA76386.1; -.
DR   EMBL; AK089211; BAC40794.1; -.
DR   EMBL; BC003892; AAH03892.1; -.
DR   EMBL; BC003904; AAH03904.1; -.
DR   PIR; A36479; A36479.
DR   HSSP; P00740; 1EDM.
DR   MGD; MGI:102768; Mfge8.
DR   GO; GO:0009897; C:external side of plasma membrane; IDA.
DR   GO; GO:0005615; C:extracellular space; IDA.
DR   GO; GO:0019897; C:extrinsic to plasma membrane; IDA.
DR   GO; GO:0005178; F:integrin binding; IDA.
DR   GO; GO:0008429; F:phosphatidylethanolamine binding; IDA.
DR   GO; GO:0001786; F:phosphatidylserine binding; IDA.
DR   GO; GO:0006910; P:phagocytosis, binding; IDA.
DR   GO; GO:0006911; P:phagocytosis, engulfment; IDA.
DR   GO; GO:0050766; P:positive regulation of phagocytosis; IDA.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001438; EGF_II.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   InterPro; IPR006210; IEGF.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   PRINTS; PR00010; EGFBLOOD.
DR   PROSITE; PS00022; EGF_1; 2.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
KW   Alternative splicing; Cell adhesion; Direct protein sequencing;
KW   EGF-like domain; Fertilization; Glycoprotein; Repeat; Signal.
FT   SIGNAL       1     22
FT   CHAIN       23    463       Lactadherin.
FT   DOMAIN      24     61       EGF-like 1.
```

```
FT   DOMAIN       64    108       EGF-like 2.
FT   DOMAIN      148    303       F5/8 type C 1.
FT   DOMAIN      308    463       F5/8 type C 2.
FT   SITE         87     89       Cell attachment site (Potential).
FT   DISULFID     28     39       By similarity.
FT   DISULFID     33     49       By similarity.
FT   DISULFID     51     60       By similarity.
FT   DISULFID     68     79       By similarity.
FT   DISULFID     73     96       By similarity.
FT   DISULFID     98    107       By similarity.
FT   DISULFID    148    303       By similarity.
FT   DISULFID    290    294       By similarity.
FT   DISULFID    308    463       By similarity.
FT   CARBOHYD     61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    266    266       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    316    316       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    426    426       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC    110    147       ETNYYNLDGEYMFTTAVPNTAVPTPAPTPDLSNNLASR ->
FT                                G (in isoform 2).
FT                                /FTId=VSP_009880.
FT   CONFLICT     35     35       N -> D (in Ref. 1; AA sequence).

  Query Match             42.4%;  Score 205;  DB 1;  Length 463;
  Best Local Similarity   42.2%;  Pred. No. 6e-14;
  Matches   38;  Conservative   16;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVF 61
              |: |||||||     :|||: ||| :       ||  : :: | || |||:: : |  |||
Db        361 NSAKEWLQVDLGTQRQVTGIITQGARDFGHIQYVASYKVAHSDDGVQWTVYEEQGSSKVF 420

Qy         62 QGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||| |: :   |  : | : ||:|: | ||
Db        421 QGNLDNNSHKKNIFEKPFMARYVRVLPVSW 450


RESULT 14
Q7TPK2
ID   Q7TPK2       PRELIMINARY;      PRT;  2102 AA.
AC   Q7TPK2;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Ac2-120.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Xu C.S., Li W.Q., Li Y.C., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA   Yang K.J., Yan H.M., Chang C.F., Zhao L.F., Ma H., Wang L., Wang S.F.,
RA   Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL   Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AY321333; AAP86265.1; -.
DR   HSSP; P12259; 1CZT.
DR   GO; GO:0007596; P:blood coagulation; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR009271; LSPD.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   Pfam; PF06049; LSPR; 20.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 1.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
SQ   SEQUENCE   2102 AA;  237257 MW;  001BCB7DE31AC560 CRC64;

  Query Match             42.1%;  Score 204;  DB 2;  Length 2102;
  Best Local Similarity   41.3%;  Pred. No. 4.3e-13;
  Matches   38;  Conservative   20;  Mismatches   32;  Indels    2;  Gaps    1;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--K 59
              || |:|||:|  |  ||| : ||| ||| : |||| : |  |  |  |  : |  :  |
Db       1995 NNNKQWLQIDLLKIKKVTAIVTQGCKSLSSEMYVKSYSILYSDQGVSWKPYRQKSSMVDK 2054

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :|:|| ::   : |  :||:::|::|| |::|
Db       2055 IFEGNSNTKGHMKNFFNPPIISRFIRIIPKTW 2086


RESULT 15
Q804X4
ID   Q804X4       PRELIMINARY;      PRT;  745 AA.
AC   Q804X4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Coagulation factor V (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA   Tuddenham E.G.D., McVey J.H.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF465271; AAO33366.1; -.
DR   HSSP; P12259; 1CZT.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_1.
FT   NON_TER       1       1
SQ   SEQUENCE   745 AA;  85626 MW;  A264587A348C29EE CRC64;

  Query Match             41.9%;  Score 203;  DB 2;  Length 745;
  Best Local Similarity   40.2%;  Pred. No. 1.7e-13;
  Matches   37;  Conservative   22;  Mismatches   31;  Indels    2;  Gaps    1;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGK--VK 59
              || ::|||:|     |:| : ||||||: |  :|| ::|  |  | :|  : ::     |
Db        636 NNNQQWLQIDLLTIKKITAIATQGVKSVTTENFVKTYVILYSNQGSEWKSYTEDSSSVAK 695

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              || ||:||   | :  :||:|:|::|| |::|
Db        696 VFSGNEDSRGHVKHFFNPPILSRFIRIVPKTW 727


Search completed: April 20, 2005, 17:21:04
Job time : 23.75 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:08:58 ; Search time 97.6188 Seconds
(without alignments)
1240.089 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332
Perfect score: 1666
Sequence: 1 LIGEHLHAGMSTLFLVYSNK.....VNSLDPPLLTRYLRIHPQSW 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1666	100.0	684	2	AAR73022	Aar73022 Human Fac
2	1666	100.0	684	2	AAR74091	Aar74091 Factor-VI
3	1666	100.0	770	3	ADP20767	Adp20767 Factor VI
4	1666	100.0	790	5	ADE64594	Ade64594 Recombina
5	1666	100.0	1014	1	AAP71139	Aap71139 Facor VII
6	1666	100.0	1383	2	AAW33227	Aaw33227 Procoagul
7	1666	100.0	1383	2	AAW33228	Aaw33228 Procoagul
8	1666	100.0	1383	2	AAW33229	Aaw33229 Procoagul
9	1666	100.0	1421	8	ADQ37613	Adq37613 Human Fac
10	1666	100.0	1421	8	ADQ37598	Adq37598 Human Fac
11	1666	100.0	1422	8	ADQ37591	Adq37591 Human Fac
12	1666	100.0	1422	8	ADQ37606	Adq37606 Human Fac
13	1666	100.0	1424	1	AAP80268	Aap80268 Modified
14	1666	100.0	1424	1	AAP91169	Aap91169 Sequence
15	1666	100.0	1424	4	AAB48842	Aab48842 Mutant ma
16	1666	100.0	1424	5	AAO18622	Aao18622 Human mat
17	1666	100.0	1425	1	AAP80267	Aap80267 Modified
18	1666	100.0	1425	8	ADQ37605	Adq37605 Human Fac
19	1666	100.0	1425	8	ADQ37590	Adq37590 Human Fac
20	1666	100.0	1428	8	ADQ37599	Adq37599 Human Fac
21	1666	100.0	1428	8	ADQ37604	Adq37604 Human Fac
22	1666	100.0	1428	8	ADQ37589	Adq37589 Human Fac
23	1666	100.0	1428	8	ADQ37614	Adq37614 Human Fac
24	1666	100.0	1434	8	ADQ37594	Adq37594 Human Fac
25	1666	100.0	1434	8	ADQ37609	Adq37609 Human Fac
26	1666	100.0	1435	8	ADQ37600	Adq37600 Human Fac
27	1666	100.0	1435	8	ADQ37615	Adq37615 Human Fac
28	1666	100.0	1437	8	ADQ37593	Adq37593 Human Fac
29	1666	100.0	1437	8	ADQ37608	Adq37608 Human Fac
30	1666	100.0	1438	3	AAB01262	Aab01262 B-domain
31	1666	100.0	1438	7	ADF31316	Adf31316 Variant o
32	1666	100.0	1438	7	ADM75988	Adm75988 Modified
33	1666	100.0	1438	7	ADM75986	Adm75986 Modified
34	1666	100.0	1438	7	ADM75989	Adm75989 Modified
35	1666	100.0	1438	7	ADM75983	Adm75983 Wild-type
36	1666	100.0	1440	2	AAR12971	Aar12971 Factor VI
37	1666	100.0	1440	8	ADQ37607	Adq37607 Human Fac
38	1666	100.0	1440	8	ADQ37592	Adq37592 Human Fac
39	1666	100.0	1441	8	ADQ37612	Adq37612 Human Fac
40	1666	100.0	1441	8	ADQ37597	Adq37597 Human Fac
41	1666	100.0	1442	8	ADQ37616	Adq37616 Human Fac
42	1666	100.0	1442	8	ADQ37601	Adq37601 Human Fac
43	1666	100.0	1444	8	ADQ37611	Adq37611 Human Fac
44	1666	100.0	1444	8	ADQ37596	Adq37596 Human Fac
45	1666	100.0	1447	5	ABG92541	Abg92541 5Arg B-do

ALIGNMENTS

RESULT 1
AAR73022
ID AAR73022 standard; peptide; 684 AA.
XX
AC AAR73022;
XX
DT 25-MAR-2003 (revised)
DT 21-NOV-1995 (first entry)
XX
DE Human Factor-VIII C-terminal fragment.
XX
KW Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
XX
OS Homo sapiens.
XX
PN WO9513301-A1.
XX
PD 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-DK000424.
XX
PR 12-NOV-1993; 93DK-00001281.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Persson E;
XX
DR WPI; 1995-194038/25.
XX
PT Crosslinked Factor VIII polypeptide which is stable - is prepd. using
PT bis(sulphosuccinimydyl) suberate or disuccinimydyl suberate in the
PT presence of polysorbate 80 to produce a coagulant with long lasting
PT activity.
XX
PS Disclosure; Page 24; 36pp; English.
XX
CC This is the C-terminal fragment (corresponding to AAs 1649-2332) of human
CC Factor-VIII which may be crosslinked resulting in increased stability and
CC retention of high activity over extended periods of time after activation
CC by thrombin. The polypeptide is used to prevent or treat diseases caused
CC by the absence or deficiency of Factor-VIII in a subject such as
CC haemophilia. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 684 AA;

Query Match 100.0%; Score 1666; DB 2; Length 684;
Best Local Similarity 100.0%; Pred. No. 2.3e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 412

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 472

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 532

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 592

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 652

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db        653 LLTRYLRIHPQSW 665
```

RESULT 2
AAR74091
ID AAR74091 standard; protein; 684 AA.
XX
AC AAR74091;
XX
DT 25-MAR-2003 (revised)
DT 04-NOV-1995 (first entry)
XX
DE Factor-VIII light chain C-terminal fragment.
XX
KW human; Factor VIII; light chain; C-terminal fragment; thrombin cleavage;
KW blood-clotting.
XX
OS Homo sapiens.
XX
PN WO9513300-A1.
XX
PD 18-MAY-1995.
XX
PF 10-NOV-1994; 94WO-DK000423.
XX
PR 12-NOV-1993; 93DK-00001280.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Kjalke M, Ezban Rasmussen M;
XX
DR WPI; 1995-194037/25.
XX
PT Factor VIII polypeptide(s) comprising a heavy chain shorter than native
PT A1-A2 domain - are easier to produce recombinantly and retain coagulant
PT activity, may be used to treat patients who have developed antibodies to
PT C-terminal epitope(s) of Factor VIII.
XX
PS Claim 4; Page 33-35; 51pp; English.
XX
CC The sequence represents C-terminal residues 1649-2332 of a human Factor-
CC VIII light chain. The sequence is produced by treating a full-length
CC Factor-VIII polypeptide with a protease, e.g. thrombin. The fragment may
CC be produced recombinantly in conjunction with a C-terminally truncated
CC heavy chain to reduce production costs and improve safety, and production
CC levels and stability are higher than for the full-length form. When the
CC fragment is used along with a C-terminally truncated heavy chain, it may
CC be used to treat patients who have developed antibodies against epitopes
CC in the C-terminal part of the heavy chain. (Updated on 25-MAR-2003 to
CC correct PN field.)

XX
SQ Sequence 684 AA;

Query Match 100.0%; Score 1666; DB 2; Length 684;
Best Local Similarity 100.0%; Pred. No. 2.3e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 412

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 472

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 532

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 592

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 652

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db        653 LLTRYLRIHPQSW 665
```

RESULT 3
ADP20767
ID ADP20767 standard; protein; 770 AA.
XX
AC ADP20767;
XX
DT 12-AUG-2004 (first entry)
XX
DE Factor VIII protein.
XX
KW A-domain; C-domain; blood coagulation; factor VIII;
KW von Willebrand factor.
XX
OS Unidentified.
XX
PN KR99066381-A.
XX
PD 16-AUG-1999.
XX
PF 24-JAN-1998; 98KR-00002264.
XX
PR 24-JAN-1998; 98KR-00002264.
XX
PA (GREC) KOREA GREEN CROSS CORP.
XX
PI Kim HC, Byun TH, Kim SM, Kim JM, Lee SG;
XX
DR WPI; 2000-547435/50.
DR N-PSDB; ADP20766.
XX
PT Cell line expressing A-domain, C-domain of blood coagulation factor VIII
PT and von Willebrand factor simultaneously, method - for manufacturing
PT recombinant blood coagulation factor VIII using the same.
XX
PS Disclosure; Fig 2; 21pp; Korean.
XX
CC The invention relates to a novel cell line expressing A-domain, C-domain
CC of blood coagulation factor VIII and von Willebrand factor
CC simultaneously, and a method for manufacturing recombinant blood
CC coagulation factor VIII using the same. The present sequence represents a
CC factor VIII protein.

```
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 1666;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-168;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        439 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 498

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        499 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 558

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        559 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 618

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        619 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 678

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        679 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 738

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db        739 LLTRYLRIHPQSW 751


RESULT 4
ADE64594
ID   ADE64594 standard; protein; 790 AA.
XX
AC   ADE64594;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Recombinant blood coagulation factor VIII protein, SEQ ID 4.
XX
KW   blood coagulation factor VIII; type-A haemophilia.
XX
OS   Unidentified.
XX
PN   CN1361178-A.
XX
PD   31-JUL-2002.
XX
PF   29-DEC-2000; 2000CN-00137779.
XX
PR   29-DEC-2000; 2000CN-00137779.
XX
PA   (SHAN-) SHANGHAI BIO-CHEM INST CHINESE ACAD SCI.
XX
PI   Qi Z,  Wang Q,  Chen C;
XX
DR   WPI; 2002-741852/81.
DR   N-PSDB; ADE64593.
XX
PT   New recombinant blood coagulation factor VIII and its production process
PT   and medicinal composition.
XX
PS   Claim 2; SEQ ID NO 4; 31pp; Chinese.
XX
CC   The invention relates to a novel recombinant blood coagulation factor
CC   VIII, its production process and its medicinal composite for treating
CC   type-A haemophilia. The invention further comprises a medicinal
CC   composition containing the blood coagulation factor which promotes blood
CC   coagulation to the blood plasma of type-A haemophilia patients. This
CC   sequence represents a recombinant blood coagulation factor VIII protein
CC   of the invention.
XX
SQ   Sequence 790 AA;

  Query Match             100.0%;  Score 1666;  DB 5;  Length 790;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-168;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 518

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 578

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 638

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 698

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 758

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db        759 LLTRYLRIHPQSW 771


RESULT 5
AAP71139
ID   AAP71139 standard; protein; 1014 AA.
XX
AC   AAP71139;
XX
DT   25-MAR-2003  (revised)
DT   14-MAY-1991  (first entry)
XX
DE   Facor VIII:c variant.
XX
KW   Factor VIII:c; variant; proteolysis; resistance;
KW   pro coagulation activity.
XX
OS   Homo sapiens.
XX
PN   WO8707144-A.
XX
PD   03-DEC-1987.
XX
PF   29-MAY-1987;   87WO-US001299.
XX
PR   29-MAY-1986;   86US-00868410.
PR   18-NOV-1986;   86US-00932767.
PR   09-DEC-1986;   86US-00939658.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Kaufman RJ,  Pittman D,  Toole JJ;
XX
DR   WPI; 1987-348539/49.
XX
PT   New deletion and replacement variants of factor 8:c - resistant to
PT   proteolysis but retaining pro-coagulant activity, and new DNA coding
PT   sequences.
XX
PS   Disclosure; Page 1; 42pp; English.
XX
CC   The full-length human factor VIII:c cDNA has been set forth in detail in
CC   WO8501961. This sequences is an example of the formula A-X-B, wherein
CC   A=Ala1-Arg372, B=Ser1690-Tyr2332 and X=0-1316 amino acids substantially
```

```
CC   duplicative of sequences of amino acids within the sequence Arg372-
CC   Ser1690 of the full-length sequence. Here X=0; producing a Arg371-Ser1690
CC   fusion protein. One or more deletions or replacements at Arg 220, 226,
CC   279, 282, 336, 359, 1719, 1721 and Lys 325, 338 and Tyr 346 will produce
CC   variants which are more resistant to specific proteolytic cleavage
CC   compared with natural factor VIII:c. Pro-coagulant activity and thrombin
CC   activatibility have been retained. See also AAP71726-29. (Updated on 25-
CC   MAR-2003 to correct PA field.)
XX
SQ   Sequence 1014 AA;

  Query Match           100.0%;  Score 1666;  DB 1;  Length 1014;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-168;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 742

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 802

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 862

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        863 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 922

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        923 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 982

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db        983 LLTRYLRIHPQSW 995


RESULT 6
AAW33227
ID   AAW33227 standard; protein; 1383 AA.
XX
AC   AAW33227;
XX
DT   21-OCT-2004  (revised)
DT   30-APR-1998  (first entry)
XX
DE   Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
KW   Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW   recombinant secretion; pro-coagulant activity; resistance;
KW   activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW   von Willebrand factor binding site; binding affinity;
KW   FVIII replacement therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .346
FT                   /note= "factor VIIIA heavy chain"
FT   Domain          1. .329
FT                   /note= "A1 domain"
FT   Domain          1. .179
FT                   /note= "plastocyanin-like domain 1"
FT   Disulfide-bond  153. .179
FT                   /note= "probable"
FT   Domain          187. .329
FT                   /note= "plastocyanin-like domain 2"
FT   Cleavage-site   372. .373
FT                   /note= "by thrombin"
FT   Domain          380. .711
FT                   /note= "A2 domain"
FT   Domain          380. .554
FT                   /note= "plastocyanin-like domain 3"
FT   Disulfide-bond  528. .554
FT                   /note= "probable"
FT   Domain          564. .711
FT                   /note= "plastocyanin-like domain 4"
FT   Region          711. .746
FT                   /note= "a spacer of the sequence
FT                   SFSQNSRHPSTRQKQFNATTIPENDIEKTDPWF AHRTPMPKIQNVSSSDLLMLL
FT                   is inserted between domains A2 and A3"
FT   Misc-difference 740
FT                   /label= R740A
FT                   /note= "wild type Arg replaced with Ala"
FT   Region          741. .1383
FT                   /note= "factor VIIIA light chain"
FT   Domain          746. .1073
FT                   /note= "A3 domain"
FT   Domain          1073. .1221
FT                   /note= "C1 domain"
FT   Domain          1226. .1378
FT                   /note= "C2 domain"
XX
PN   WO9740145-A1.
XX
PD   30-OCT-1997.
XX
PF   24-APR-1997;   97WO-US006563.
XX
PR   24-APR-1996;   96US-0016117P.
PR   15-MAY-1996;   96US-0017785P.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Kaufman RJ,  Pipe SW,  Amano K;
XX
DR   WPI; 1997-535830/49.
XX
PT   Modified human pro-coagulant active factor VIII - can be administered to
PT   haemophiliacs, i.e. factor VIII replacement therapy.
XX
PS   Claim 20; Page; 57pp; English.
XX
CC   The present sequence represents a novel pro-coagulant active factor VIII
CC   (FVIII) mutant protein, comprising a deletion of the B domain and von
CC   Willebrand factor binding site, a mutation at Arg740 and an addition of
CC   an amino acid sequence spacer between the A2 and A3 domains. Factor VIII,
CC   along with calcium and phospholipid, acts as a cofactor for factor IXA,
CC   when it converts factor X to the activated form (factor XA). FVIII is the
CC   coagultion factor deficient in the X-chromosome-linked bleeding disorder
CC   haemophilia A. Several other mutant FVIII proteins have also been created
CC   (see AAW33222-29). The FVIII mutant F309S (AAW33225) is capable of
CC   recombinant secretion at higher levels than typically obtained with wild
CC   type FVIII and retains pro-coagulant activity. The FVIII mutant R336I
CC   (AAW33222) and R562K (AAW33223) are resistant to activated protein C
CC   (APC) cleavage. The present FVIII mutant can form a more stable
CC   configuration, and have an approximate 5-fold increase in specific
CC   activity compared to purified wild type FVIII, while increasing their
CC   binding affinity to von Willebrand factor improves their stability. The
CC   FVIII proteins can be administered to haemophiliacs, i.e. FVIII
CC   replacement therapy, while the nucleic acid molecule can be used for gene
CC   therapy. note: this sequence does not appear in the specification; it was
CC   created using sequences from the given references
CC
CC   Revised record issued on 21-OCT-2004 : Correction to feature table key
XX
SQ   Sequence 1383 AA;

  Query Match           100.0%;  Score 1666;  DB 2;  Length 1383;
  Best Local Similarity  100.0%;  Pred. No. 6.7e-168;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1052 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1111

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1112 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1171

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1172 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1231

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1232 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1291

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1292 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1351

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1352 LLTRYLRIHPQSW 1364
```

RESULT 7
AAW33228
ID AAW33228 standard; protein; 1383 AA.
XX
AC AAW33228;
XX
DT 21-OCT-2004 (revised)
DT 30-APR-1998 (first entry)
XX
DE Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
KW Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW recombinant secretion; pro-coagulant activity; resistance;
KW activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW von Willebrand factor binding site; binding affinity;
KW FVIII replacement therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX

FH	Key	Location/Qualifiers
FT	Region	1. .346
FT		/note= "factor VIIIA heavy chain"
FT	Domain	1. .329
FT		/note= "A1 domain"
FT	Domain	1. .179
FT		/note= "plastocyanin-like domain 1"
FT	Disulfide-bond	153. .179
FT		/note= "probable"
FT	Domain	187. .329
FT		/note= "plastocyanin-like domain 2"
FT	Misc-difference	336
FT		/label= R336I
FT		/note= "wild type Arg replaced with Ile"
FT	Cleavage-site	372. .373
FT		/note= "by thrombin"
FT	Domain	380. .711
FT		/note= "A2 domain"
FT	Domain	380. .554
FT		/note= "plastocyanin-like domain 3"
FT	Disulfide-bond	528. .554
FT		/note= "probable"
FT	Misc-difference	562
FT		/label= R562K
FT		/note= "wild type Arg replaced with Lys"
FT	Domain	564. .711
FT		/note= "plastocyanin-like domain 4"
FT	Region	711. .746
FT		/note= "a spacer of the sequence SFSQNSRHPSTRQKQFNATTIPENDIEKTDPWF AHRTPMPKIQNVSSSDLLMLL is inserted between domains A2 and A3"
FT	Misc-difference	740
FT		/label= R740A
FT		/note= "wild type Arg replaced with Ala"
FT	Region	741. .1383
FT		/note= "factor VIIIA light chain"
FT	Domain	746. .1073
FT		/note= "A3 domain"
FT	Domain	1073. .1221
FT		/note= "C1 domain"
FT	Domain	1226. .1378
FT		/note= "C2 domain"

XX
PN WO9740145-A1.
XX
PD 30-OCT-1997.
XX
PF 24-APR-1997; 97WO-US006563.
XX
PR 24-APR-1996; 96US-0016117P.
PR 15-MAY-1996; 96US-0017785P.
XX
PA (UNMI) UNIV MICHIGAN.
XX
PI Kaufman RJ, . Pipe SW, Amano K;
XX
DR WPI; 1997-535830/49.
XX
PT Modified human pro-coagulant active factor VIII - can be administered to
PT haemophiliacs, i.e. factor VIII replacement therapy.
XX
PS Claim 18; Page; 57pp; English.
XX
CC The present sequence represents a novel pro-coagulant active factor VIII
CC (FVIII) mutant protein, comprising a deletion of the B domain and von
CC Willebrand factor binding site, mutations R336I, R562K and R740A and an
CC addition of an amino acid sequence spacer between the A2 and A3 domains.
CC Factor VIII, along with calcium and phospholipid, acts as a cofactor for
CC factor IXA, when it converts factor X to the activated form (factor XA).
CC FVIII is the coagultion factor deficient in the X-chromosome-linked
CC bleeding disorder haemophilia A. Several other mutant FVIII proteins have
CC also been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is
CC capable of recombinant secretion at higher levels than typically obtained
CC with wild type FVIII and retains pro-coagulant activity. The FVIII mutant
CC R336I (AAW33222) and R562K (AAW33223) are resistant to activated protein
CC C (APC) cleavage. The FVIII mutant comprising a deletion of the B domain
CC and von Willebrand factor binding site, a mutation at Arg740 and an
CC addition of an amino acid sequence spacer between the A2 and A3 domains
CC can form a more stable configuration, and have an approximate 5-fold
CC increase in specific activity compared to purified wild type FVIII, while
CC increasing their binding affinity to von Willebrand factor improves their
CC stability. The FVIII proteins can be administered to haemophiliacs, i.e.
CC FVIII replacement therapy, while the nucleic acid molecule can be used
CC for gene therapy. note: this sequence does not appear in the
CC specification; it was created using sequences from the given references
CC
CC Revised record issued on 21-OCT-2004 : Correction to feature table key
XX
SQ Sequence 1383 AA;

Query Match 100.0%; Score 1666; DB 2; Length 1383;
Best Local Similarity 100.0%; Pred. No. 6.7e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1052 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1111

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1112 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1171
```

```
Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1172 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1231

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1232 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1291

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1292 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1351

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1352 LLTRYLRIHPQSW 1364
```

```
RESULT 8
AAW33229
ID   AAW33229 standard; protein; 1383 AA.
XX
AC   AAW33229;
XX
DT   21-OCT-2004  (revised)
DT   30-APR-1998  (first entry)
XX
DE   Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
KW   Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW   recombinant secretion; pro-coagulant activity; resistance;
KW   activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW   von Willebrand factor binding site; binding affinity;
KW   FVIII replacement therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .346
FT                   /note= "factor VIIIA heavy chain"
FT   Domain          1. .329
FT                   /note= "A1 domain"
FT   Domain          1. .179
FT                   /note= "plastocyanin-like domain 1"
FT   Disulfide-bond  153. .179
FT                   /note= "probable"
FT   Domain          187. .329
FT                   /note= "plastocyanin-like domain 2"
FT   Misc-difference 309
FT                   /label= F309S
FT                   /note= "wild type Phe replaced with Ser"
FT   Cleavage-site   372. .373
FT                   /note= "by thrombin"
FT   Domain          380. .711
FT                   /note= "A2 domain"
FT   Domain          380. .554
FT                   /note= "plastocyanin-like domain 3"
FT   Disulfide-bond  528. .554
FT                   /note= "probable"
FT   Domain          564. .711
FT                   /note= "plastocyanin-like domain 4"
FT   Region          711. .746
FT                   /note= "a spacer of the sequence
FT                   SFSQNSRHPSTRQKQFNATTIPENDIEKTDPWF AHRTPMPKIQNVSSSDLLMLL
FT                   is inserted between domains A2 and A3"
FT   Misc-difference 740
FT                   /label= R740A
FT                   /note= "wild type Arg replaced with Ala"
FT   Region          741. .1383
FT                   /note= "factor VIIIA light chain"
FT   Domain          746. .1073
FT                   /note= "A3 domain"
FT   Domain          1073. .1221
FT                   /note= "C1 domain"
FT   Domain          1226. .1378
FT                   /note= "C2 domain"
XX
PN   WO9740145-A1.
XX
PD   30-OCT-1997.
XX
PF   24-APR-1997;   97WO-US006563.
XX
PR   24-APR-1996;   96US-0016117P.
PR   15-MAY-1996;   96US-0017785P.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Kaufman RJ,  Pipe SW,  Amano K;
XX
DR   WPI; 1997-535830/49.
XX
PT   Modified human pro-coagulant active factor VIII - can be administered to
PT   haemophiliacs, i.e. factor VIII replacement therapy.
XX
PS   Claim 19; Page; 57pp; English.
XX
CC   The present sequence represents a novel pro-coagulant active factor VIII
CC   (FVIII) mutant protein, comprising a deletion of the B domain and von
CC   Willebrand factor binding site, mutations F309S, R740A and addition of an
CC   amino acid sequence spacer between the A2 and A3 domains. Factor VIII,
CC   along with calcium and phospholipid, acts as a cofactor for factor IXA,
CC   when it converts factor X to the activated form (factor XA). FVIII is the
CC   coagultion factor deficient in the X-chromosome-linked bleeding disorder
CC   haemophilia A. Several other mutant FVIII proteins have also been created
CC   (see AAW33222-29). The FVIII mutant F309S (AAW33225) is capable of
CC   recombinant secretion at higher levels than typically obtained with wild
CC   type FVIII and retains pro-coagulant activity. The FVIII mutant R336I
CC   (AAW33222) and R562K (AAW33223) are resistant to activated protein C
CC   (APC) cleavage. The present FVIII mutant can form a more stable
CC   configuration, and have an approximate 5-fold increase in specific
CC   activity compared to purified wild type FVIII, while increasing their
CC   binding affinity to von Willebrand factor improves their stability. The
CC   FVIII proteins can be administered to haemophiliacs, i.e. FVIII
CC   replacement therapy, while the nucleic acid molecule can be used for gene
CC   therapy. note: this sequence does not appear in the specification; it was
CC   created using sequences from the given references
CC
CC   Revised record issued on 21-OCT-2004 : Correction to feature table key
XX
SQ   Sequence 1383 AA;
```

```
  Query Match           100.0%;  Score 1666;  DB 2;  Length 1383;
  Best Local Similarity  100.0%;  Pred. No. 6.7e-168;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1052 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1111

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1112 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1171

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1172 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1231

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1232 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1291

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1292 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1351
```

```
Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1352 LLTRYLRIHPQSW 1364


RESULT 9
ADQ37613
ID   ADQ37613 standard; protein; 1421 AA.
XX
AC   ADQ37613;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Human Factor VIII deletion Pro739Phe mutant dB747.
XX
KW   Human; Factor VIII; clotting factor; blood; blood disorder;
KW   haemophilia A; gene therapy; mutant; mutein; B domain.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 739
FT                   /note= "Wild-type Pro substituted by Phe"
FT   Misc-difference 747. .748
FT                   /note= "Amino acids 748-1658 of the wild-type sequence
FT                   have been deleted"
XX
PN   US2004147436-A1.
XX
PD   29-JUL-2004.
XX
PF   28-JAN-2003; 2003US-00353753.
XX
PR   28-JAN-2003; 2003US-00353753.
XX
PA   (KIMH/) KIM H.
PA   (SONG/) SONG I.
PA   (CHOI/) CHOI J W.
PA   (JANG/) JANG J.
PA   (KIMY/) KIM Y.
PA   (LEEH/) LEE H S.
PA   (BANG/) BANG Y.
PA   (KIMD/) KIM D.
XX
PI   Kim H,  Song I,  Choi JW,  Jang J,  Kim Y,  Lee HS,  Bang Y,  Kim D;
XX
DR   WPI; 2004-561406/54.
XX
PT   New Factor VIII polypeptides that are more stable than full-length Factor
PT   VIII, useful for clotting blood or for preventing or treating blood
PT   disorders, such as Hemophilia A.
XX
PS   Example 4; Page; 39pp; English.
XX
CC   The invention relates to a Factor VIII polypeptide comprising an internal
CC   deletion of one or more amino acids between 1649 and 1688 fused to any
CC   amino acid sequence in B domain from about 741-782, with reference to
CC   full-length human Factor VIII appearing as ADQ37564. Also included are a
CC   pharmaceutical composition comprising the Factor VIII polypeptide (and a
CC   pharmaceutical carrier), a lyophilised composition comprising the Factor
CC   VIII polypeptide, clotting blood in a subject (comprising contacting a
CC   clotting amount of the Factor VIII polypeptide with the blood), treating
CC   Haemophilia A in a patient (comprising administering a clotting amount of
CC   the above Factor VIII polypeptide to a patient), an isolated nucleic acid
CC   encoding the Factor VIII polypeptide, an expression vector comprising the
CC   nucleic acid operably linked to a promoter, a host cell comprising the
CC   expression vector, making the Factor VIII polypeptide (comprising
CC   culturing the cell in conditions suitable for the vector to express the
CC   polypeptide, and isolating the polypeptide) and a purified antibody
CC   specific for the Factor VIII polypeptide cited above. The internal
CC   deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
CC   758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
CC   -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single
CC   chain. The proline at 739 is replaced by another amino acid. The
CC   polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
CC   encompassing fusion sites between Asn amino acid at positions 745,757 or
CC   764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
CC   reference to the full-length human Factor VIII amino acid sequence cited
CC   above. The polypeptide comprises a formula with the following linked
CC   domains: H-S-L, where the H domain represents a polypeptide sequence
CC   comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
CC   the S domain represents a polypeptide spacing linker comprising up to
CC   about 60 amino acids and the L domain represents a polypeptide sequence
CC   comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
CC   composition and methods are useful for clotting blood or for preventing
CC   or treating blood disorders, such as Haemophilia A. The present sequence
CC   is a human factor VIII B domain deletion mutant also carrying the
CC   Pro739Phe mutation. NOTE: The present sequence is not shown in the
CC   specification but was created by the indexer using information in the
CC   specification.
XX
SQ   Sequence 1421 AA;

  Query Match             100.0%;  Score 1666;  DB 8;  Length 1421;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-168;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1149

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1209

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1269

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1329

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1389

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1390 LLTRYLRIHPQSW 1402


RESULT 10
ADQ37598
ID   ADQ37598 standard; protein; 1421 AA.
XX
AC   ADQ37598;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Human Factor VIII deletion mutant dB747.
XX
KW   Human; Factor VIII; clotting factor; blood; blood disorder;
KW   haemophilia A; gene therapy; mutant; mutein; B domain.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 747. .748
FT                   /note= "Amino acids 748-1658 of the wild-type sequence
FT                   have been deleted"
XX
PN   US2004147436-A1.
```

```
XX
PD   29-JUL-2004.
XX
PF   28-JAN-2003; 2003US-00353753.
XX
PR   28-JAN-2003; 2003US-00353753.
XX
PA   (KIMH/) KIM H.
PA   (SONG/) SONG I.
PA   (CHOI/) CHOI J W.
PA   (JANG/) JANG J.
PA   (KIMY/) KIM Y.
PA   (LEEH/) LEE H S.
PA   (BANG/) BANG Y.
PA   (KIMD/) KIM D.
XX
PI   Kim H,  Song I,  Choi JW,  Jang J,  Kim Y,  Lee HS,  Bang Y,  Kim D;
XX
DR   WPI; 2004-561406/54.
XX
PT   New Factor VIII polypeptides that are more stable than full-length Factor
PT   VIII, useful for clotting blood or for preventing or treating blood
PT   disorders, such as Hemophilia A.
XX
PS   Claim 2; Page; 39pp; English.
XX
CC   The invention relates to a Factor VIII polypeptide comprising an internal
CC   deletion of one or more amino acids between 1649 and 1688 fused to any
CC   amino acid sequence in B domain from about 741-782, with reference to
CC   full-length human Factor VIII appearing as ADQ37564. Also included are a
CC   pharmaceutical composition comprising the Factor VIII polypeptide (and a
CC   pharmaceutical carrier), a lyophilised composition comprising the Factor
CC   VIII polypeptide, clotting blood in a subject (comprising contacting a
CC   clotting amount of the Factor VIII polypeptide with the blood), treating
CC   Haemophilia A in a patient (comprising administering a clotting amount of
CC   the above Factor VIII polypeptide to a patient), an isolated nucleic acid
CC   encoding the Factor VIII polypeptide, an expression vector comprising the
CC   nucleic acid operably linked to a promoter, a host cell comprising the
CC   expression vector, making the Factor VIII polypeptide (comprising
CC   culturing the cell in conditions suitable for the vector to express the
CC   polypeptide, and isolating the polypeptide) and a purified antibody
CC   specific for the Factor VIII polypeptide cited above. The internal
CC   deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
CC   758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
CC   -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single
CC   chain. The proline at 739 is replaced by another amino acid. The
CC   polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
CC   encompassing fusion sites between Asn amino acid at positions 745,757 or
CC   764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
CC   reference to the full-length human Factor VIII amino acid sequence cited
CC   above. The polypeptide comprises a formula with the following linked
CC   domains: H-S-L, where the H domain represents a polypeptide sequence
CC   comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
CC   the S domain represents a polypeptide spacing linker comprising up to
CC   about 60 amino acids and the L domain represents a polypeptide sequence
CC   comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
CC   composition and methods are useful for clotting blood or for preventing
CC   or treating blood disorders, such as Haemophilia A. The present sequence
CC   is a human factor VIII B domain deletion mutant. NOTE: The present
CC   sequence is not shown in the specification but was created by the indexer
CC   using information in the specification.
XX
SQ   Sequence 1421 AA;

  Query Match             100.0%;  Score 1666;  DB 8;  Length 1421;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-168;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1090 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1149

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1150 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1209

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1210 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1269

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1270 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1329

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1389

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1390 LLTRYLRIHPQSW 1402


RESULT 11
ADQ37591
ID   ADQ37591 standard; protein; 1422 AA.
XX
AC   ADQ37591;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Human Factor VIII Glycosylation/deletion mutant dBN(45-56).
XX
KW   Human; Factor VIII; clotting factor; blood; blood disorder;
KW   haemophilia A; gene therapy; mutant; mutein; N-glycosylation.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   745. .747
FT                   /note= "Asn is N-glycosylated"
FT   Misc-difference 745. .746
FT                   /note= "Amino acids 746-1655 of the wild-type sequence
FT                   have been deleted"
XX
PN   US2004147436-A1.
XX
PD   29-JUL-2004.
XX
PF   28-JAN-2003; 2003US-00353753.
XX
PR   28-JAN-2003; 2003US-00353753.
XX
PA   (KIMH/) KIM H.
PA   (SONG/) SONG I.
PA   (CHOI/) CHOI J W.
PA   (JANG/) JANG J.
PA   (KIMY/) KIM Y.
PA   (LEEH/) LEE H S.
PA   (BANG/) BANG Y.
PA   (KIMD/) KIM D.
XX
PI   Kim H,  Song I,  Choi JW,  Jang J,  Kim Y,  Lee HS,  Bang Y,  Kim D;
XX
DR   WPI; 2004-561406/54.
XX
PT   New Factor VIII polypeptides that are more stable than full-length Factor
PT   VIII, useful for clotting blood or for preventing or treating blood
PT   disorders, such as Hemophilia A.
XX
PS   Claim 2; Page; 39pp; English.
XX
CC   The invention relates to a Factor VIII polypeptide comprising an internal
CC   deletion of one or more amino acids between 1649 and 1688 fused to any
CC   amino acid sequence in B domain from about 741-782, with reference to
```

CC full-length human Factor VIII appearing as ADQ37564. Also included are a
CC pharmaceutical composition comprising the Factor VIII polypeptide (and a
CC pharmaceutical carrier), a lyophilised composition comprising the Factor
CC VIII polypeptide, clotting blood in a subject (comprising contacting a
CC clotting amount of the Factor VIII polypeptide with the blood), treating
CC Haemophilia A in a patient (comprising administering a clotting amount of
CC the above Factor VIII polypeptide to a patient), an isolated nucleic acid
CC encoding the Factor VIII polypeptide, an expression vector comprising the
CC nucleic acid operably linked to a promoter, a host cell comprising the
CC expression vector, making the Factor VIII polypeptide (comprising
CC culturing the cell in conditions suitable for the vector to express the
CC polypeptide, and isolating the polypeptide) and a purified antibody
CC specific for the Factor VIII polypeptide cited above. The internal
CC deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
CC 758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
CC -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single
CC chain. The proline at 739 is replaced by another amino acid. The
CC polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
CC encompassing fusion sites between Asn amino acid at positions 745,757 or
CC 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
CC reference to the full-length human Factor VIII amino acid sequence cited
CC above. The polypeptide comprises a formula with the following linked
CC domains: H-S-L, where the H domain represents a polypeptide sequence
CC comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
CC the S domain represents a polypeptide spacing linker comprising up to
CC about 60 amino acids and the L domain represents a polypeptide sequence
CC comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
CC composition and methods are useful for clotting blood or for preventing
CC or treating blood disorders, such as Haemophilia A. The present sequence
CC is a human factor VIII deletion mutant where the deletion creates an N-
CC glycosylation site. NOTE: The present sequence is not shown in the
CC specification but was created by the indexer using information in the
CC specification.
XX
SQ Sequence 1422 AA;

Query Match 100.0%; Score 1666; DB 8; Length 1422;
Best Local Similarity 100.0%; Pred. No. 7e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1150

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1151 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1210

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1270

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1330

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1331 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1390

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1391 LLTRYLRIHPQSW 1403
```

RESULT 12
ADQ37606
ID ADQ37606 standard; protein; 1422 AA.
XX
AC ADQ37606;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human Factor VIII Glycosylation/deletion Pro739Phe mutant dBN(45-56).
XX
KW Human; Factor VIII; clotting factor; blood; blood disorder;
KW haemophilia A; gene therapy; mutant; mutein; N-glycosylation.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 739
FT /note= "Wild-type Pro substituted by Phe"
FT Modified-site 745. .747
FT /note= "Asn is N-glycosylated"
FT Misc-difference 745. .746
FT /note= "Amino acids 746-1655 of the wild-type sequence
FT have been deleted"
XX
PN US2004147436-A1.
XX
PD 29-JUL-2004.
XX
PF 28-JAN-2003; 2003US-00353753.
XX
PR 28-JAN-2003; 2003US-00353753.
XX
PA (KIMH/) KIM H.
PA (SONG/) SONG I.
PA (CHOI/) CHOI J W.
PA (JANG/) JANG J.
PA (KIMY/) KIM Y.
PA (LEEH/) LEE H S.
PA (BANG/) BANG Y.
PA (KIMD/) KIM D.
XX
PI Kim H, Song I, Choi JW, Jang J, Kim Y, Lee HS, Bang Y, Kim D;
XX
DR WPI; 2004-561406/54.
XX
PT New Factor VIII polypeptides that are more stable than full-length Factor
PT VIII, useful for clotting blood or for preventing or treating blood
PT disorders, such as Hemophilia A.
XX
PS Example 4; Page; 39pp; English.
XX
CC The invention relates to a Factor VIII polypeptide comprising an internal
CC deletion of one or more amino acids between 1649 and 1688 fused to any
CC amino acid sequence in B domain from about 741-782, with reference to
CC full-length human Factor VIII appearing as ADQ37564. Also included are a
CC pharmaceutical composition comprising the Factor VIII polypeptide (and a
CC pharmaceutical carrier), a lyophilised composition comprising the Factor
CC VIII polypeptide, clotting blood in a subject (comprising contacting a
CC clotting amount of the Factor VIII polypeptide with the blood), treating
CC Haemophilia A in a patient (comprising administering a clotting amount of
CC the above Factor VIII polypeptide to a patient), an isolated nucleic acid
CC encoding the Factor VIII polypeptide, an expression vector comprising the
CC nucleic acid operably linked to a promoter, a host cell comprising the
CC expression vector, making the Factor VIII polypeptide (comprising
CC culturing the cell in conditions suitable for the vector to express the
CC polypeptide, and isolating the polypeptide) and a purified antibody
CC specific for the Factor VIII polypeptide cited above. The internal
CC deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
CC 758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
CC -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single
CC chain. The proline at 739 is replaced by another amino acid. The
CC polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
CC encompassing fusion sites between Asn amino acid at positions 745,757 or
CC 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
CC reference to the full-length human Factor VIII amino acid sequence cited
CC above. The polypeptide comprises a formula with the following linked
CC domains: H-S-L, where the H domain represents a polypeptide sequence
CC comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
CC the S domain represents a polypeptide spacing linker comprising up to
CC about 60 amino acids and the L domain represents a polypeptide sequence

CC comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
CC composition and methods are useful for clotting blood or for preventing
CC or treating blood disorders, such as Haemophilia A. The present sequence
CC is a human factor VIII deletion mutant where the deletion creates an N-
CC glycosylation site and carries the Pro739Phe mutation. NOTE: The present
CC sequence is not shown in the specification but was created by the indexer
CC using information in the specification.
XX
SQ Sequence 1422 AA;

Query Match 100.0%; Score 1666; DB 8; Length 1422;
Best Local Similarity 100.0%; Pred. No. 7e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1150

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1151 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1210

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1270

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1330

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1331 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1390

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1391 LLTRYLRIHPQSW 1403
```

RESULT 13
AAP80268
ID AAP80268 standard; protein; 1424 AA.
XX
AC AAP80268;
XX
DT 25-MAR-2003 (revised)
DT 10-OCT-1990 (first entry)
XX
DE Modified factor VIII:C sequence with the R740-D1658 deletion.
XX
KW Modified factor VIII:C; maturation polypeptide; haemophilia;
KW blood coagulation; RD deletion; procoagulant.
XX
OS Homo sapiens.
XX
PN WO8800831-A.
XX
PD 11-FEB-1988.
XX
PF 31-JUL-1987; 87WO-US001814.
XX
PR 01-AUG-1986; 86US-00893375.
XX
PA (BIOJ) BIOGEN NV.
PA (PASE/) PASEK M P.
XX
PI Pasek MP;
XX
DR WPI; 1988-049866/07.
DR N-PSDB; AAN80447.
XX
PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA
PT encoding maturation polypeptide, useful for high yield transformation.
XX
PS Claim 3; Page 60-61-62-63; 97pp; English.
XX
CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part of
CC the sequence encoding the maturation polypeptide of factor VIII:C is
CC deleted, i.e. Gln 744 - Asp 1563. The full lenght Factor VIII:C cDNA has
CC two changes with respect to the published sequence (EPO application
CC 160457): CTG to CTA at Leu 242 and TTC to CTC change at amino acid
CC residue 1880 (Phe to Leu). The product is produced in approx. 20 times
CC higher yields than previous recombinant produced factor VIII:C and are
CC more easily purified. The peptide is used for treating haemophilia A,
CC both acute and prolonged bleeding. See also AAN80444 and AAN80446.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1424 AA;

Query Match 100.0%; Score 1666; DB 1; Length 1424;
Best Local Similarity 100.0%; Pred. No. 7e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1093 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1152

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1153 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1212

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1213 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1272

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1273 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1332

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1333 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1392

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1393 LLTRYLRIHPQSW 1405
```

RESULT 14
AAP91169
ID AAP91169 standard; protein; 1424 AA.
XX
AC AAP91169;
XX
DT 25-MAR-2003 (revised)
DT 26-JUN-1990 (first entry)
XX
DE Sequence of 740 Arg-1649 Glu human Factor VIII:C.
XX
KW Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;
KW haemophilia A.
XX
OS Homo sapiens.
XX
PN EP306968-A.
XX
PD 15-MAR-1989.
XX
PF 09-SEP-1988; 88EP-00114769.
XX
PR 10-SEP-1987; 87JP-00225147.
PR 08-APR-1988; 88JP-00085454.
XX
PA (KAGA) CHEMO SERO THERAPEUTIC RES INST.
PA (TEIJ) TEIJIN LTD.
XX

```
PI   Sugiyama T,  Masuda K,  Tajima Y,  Yonemura H;
XX
DR   WPI; 1989-078467/11.
DR   N-PSDB; AAN90654.
XX
PT   Prodn. of recombinant human Factor-VIII-C - using animal cells
PT   transformed with a vector contg. the gene for Factor VIII:C and a
PT   promoter.
XX
PS   Disclosure; Fig 1; 32pp; English.
XX
CC   Arg-740 of the carboxyl terminus of the H chain is directly bonded by a
CC   peptide bond to Glu-1649 of the amino terminus of L chain. A prefd.
CC   expression vector used to transform animal cell so that they produce
CC   human Factor VIII:Cis plasmid Ad.RE.neo. The expression vector has at
CC   least one promoter upstream of AAN90654. The transformants can constantly
CC   and continuously produce human Factor VIII:C in high yield on a
CC   commercial scale. The human Factor VIII:C so produced is considered to
CC   corresp. to the smallest species of active and intact Factor VIII:C
CC   molecules in the human blood plasma. It is useful for treating
CC   haemophilia A patients. (Updated on 25-MAR-2003 to correct PR field.)
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 1424 AA;

  Query Match             100.0%;  Score 1666;  DB 1;  Length 1424;
  Best Local Similarity   100.0%;  Pred. No. 7e-168;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1093 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1152

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1153 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1212

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1213 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1272

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1273 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1332

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1333 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1392

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1393 LLTRYLRIHPQSW 1405


RESULT 15
AAB48842
ID   AAB48842 standard; protein; 1424 AA.
XX
AC   AAB48842;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Mutant mature human factor VIII, SEQ ID NO:5.
XX
KW   Factor VIII; human; B domain; LRP-mediated plasma clearance;
KW   receptor-dependent clearance; receptor-independent clearance; half-life;
KW   haemophilia; mutant; mutein.
XX
OS   Homo sapiens.
XX
PN   WO200071714-A2.
XX
PD   30-NOV-2000.
XX
PF   24-MAY-2000; 2000WO-US014111.
XX
PR   24-MAY-1999;   99US-0135847P.
XX
PA   (AMNA-) AMERICAN NAT RED CROSS.
XX
PI   Saenko EL,  Strickland DK;
XX
DR   WPI; 2001-025163/03.
XX
PT   Factor VIII mutants having increased half-life useful for treating
PT   hemophilia, comprise one or more amino acid substitutions in the A2
PT   and/or C2 domain of factor VIII.
XX
PS   Claim 9; Fig 2A-B; 121pp; English.
XX
CC   The invention relates to human factor VIII mutants comprising an amino
CC   acid substitution at one or more positions in the A2 domain and/or an
CC   amino acid substitution at one or more positions in the C2 domain. The
CC   invention also encompasses a factor VIII mutant which lacks a B domain
CC   (AAB48842). The factor VIII mutants have an increased half-life in the
CC   bloodstream. The A2 domain mutants exhibit reduced LRP-dependent
CC   (receptor-dependent) clearance of factor VIII, while C2 domain mutants
CC   have reduced receptor-independent clearance. The invention also relates
CC   to a method of using RAP (receptor associated protein), a protein which
CC   inhibits LRP (low density lipoprotein related protein)-mediated ligand
CC   internalisation, to increase the half-life of factor VIII. The mutant
CC   factor VIII proteins, and nucleotides encoding them, are useful for
CC   treating haemophilia. RAP, LRP-binding RAP mutants or fragments, and
CC   nucleic acids encoding them may also be used in the treatment of
CC   haemophilia, in combination with a mutant factor VIII protein or DNA of
CC   the invention. The invention provides means of increasing the half-life
CC   of factor VIII by reducing its clearance from plasma. The present
CC   sequence represents a mutant mature human factor VIII which lacks a B
CC   domain
XX
SQ   Sequence 1424 AA;

  Query Match             100.0%;  Score 1666;  DB 4;  Length 1424;
  Best Local Similarity   100.0%;  Pred. No. 7e-168;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1093 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1152

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1153 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1212

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1213 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1272

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1273 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1332

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1333 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1392

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1393 LLTRYLRIHPQSW 1405


Search completed: April 20, 2005, 17:19:11
Job time : 99.6188 secs
```

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:14:28 ; Search time 30.2153 Seconds
(without alignments)
773.288 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332
Perfect score: 1666
Sequence: 1 LIGEHLHAGMSTLFLVYSNK.....VNSLDPPLLTRYLRIHPQSW 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1666	100.0	1438	3	US-09-209-916-1	Sequence 1, Appli
2	1666	100.0	1457	4	US-09-001-039B-47	Sequence 47, Appl
3	1666	100.0	1471	1	US-08-683-839B-3	Sequence 3, Appli
4	1666	100.0	1661	2	US-08-882-083-2	Sequence 2, Appli
5	1666	100.0	1661	2	US-08-558-107-2	Sequence 2, Appli
6	1666	100.0	1661	3	US-09-243-539-2	Sequence 2, Appli
7	1666	100.0	2332	1	US-07-864-004B-4	Sequence 4, Appli
8	1666	100.0	2332	1	US-08-251-937A-4	Sequence 4, Appli
9	1666	100.0	2332	1	US-08-212-133A-2	Sequence 2, Appli
10	1666	100.0	2332	1	US-08-474-503-2	Sequence 2, Appli
11	1666	100.0	2332	2	US-08-670-707A-2	Sequence 2, Appli
12	1666	100.0	2332	3	US-09-037-601-2	Sequence 2, Appli
13	1666	100.0	2332	3	US-09-324-867-3	Sequence 3, Appli
14	1666	100.0	2332	3	US-09-315-179-2	Sequence 2, Appli
15	1666	100.0	2332	4	US-09-523-656-2	Sequence 2, Appli
16	1666	100.0	2332	4	US-09-957-641A-2	Sequence 2, Appli
17	1666	100.0	2332	5	PCT-US93-03275-4	Sequence 4, Appli
18	1666	100.0	2332	5	PCT-US94-13200-2	Sequence 2, Appli
19	1666	100.0	2351	1	US-08-121-202-2	Sequence 2, Appli
20	1666	100.0	2351	1	US-08-366-851A-2	Sequence 2, Appli
21	1666	100.0	2351	4	US-10-133-907-4	Sequence 4, Appli
22	1666	100.0	2351	4	US-09-001-039B-45	Sequence 45, Appl
23	1663	99.8	2332	1	US-08-276-594A-2	Sequence 2, Appli
24	1660	99.6	2351	6	5171844-2	Patent No. 5171844
25	1660	99.6	2351	6	5171844-2	Patent No. 5171844
26	1657	99.5	2351	6	5422260-1	Patent No. 5422260
27	1657	99.5	2351	6	5422260-1	Patent No. 5422260
28	1522	91.4	2343	3	US-09-324-867-2	Sequence 2, Appli
29	1495	89.7	2319	1	US-08-212-133A-8	Sequence 8, Appli
30	1495	89.7	2319	1	US-08-474-503-6	Sequence 6, Appli
31	1495	89.7	2319	2	US-08-670-707A-6	Sequence 6, Appli
32	1495	89.7	2319	3	US-09-037-601-6	Sequence 6, Appli
33	1495	89.7	2319	3	US-09-315-179-6	Sequence 6, Appli
34	1495	89.7	2319	4	US-09-523-656-28	Sequence 28, Appl
35	1495	89.7	2319	5	PCT-US94-13200-6	Sequence 6, Appli
36	1480.5	88.9	2304	3	US-09-324-867-4	Sequence 4, Appli
37	1467	88.1	1443	2	US-08-670-707A-39	Sequence 39, Appl
38	1467	88.1	1443	3	US-09-037-601-39	Sequence 39, Appl
39	1467	88.1	1443	3	US-09-315-179-39	Sequence 39, Appl
40	1467	88.1	1467	4	US-09-523-656-38	Sequence 38, Appl
41	1467	88.1	2133	2	US-08-670-707A-37	Sequence 37, Appl
42	1467	88.1	2133	3	US-09-037-601-37	Sequence 37, Appl
43	1467	88.1	2133	3	US-09-315-179-37	Sequence 37, Appl
44	1467	88.1	2133	4	US-09-523-656-30	Sequence 30, Appl
45	1454	87.3	2115	3	US-09-324-867-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-209-916-1
; Sequence 1, Application US/09209916
; Patent No. 6358703
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/09/209,916
; CURRENT FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-09-209-916-1

Query Match 100.0%; Score 1666; DB 3; Length 1438;
Best Local Similarity 100.0%; Pred. No. 9.8e-168;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1166

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1226

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1286

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1346

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1406

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
```

```
Db       1407 LLTRYLRIHPQSW 1419


RESULT 2
US-09-001-039B-47
; Sequence 47, Application US/09001039B
; Patent No. 6818439
;  GENERAL INFORMATION:
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Respess, James G.
;    APPLICANT:  DePolo, Nicholas J.
;    APPLICANT:  Hsu, David Chi-Tang
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Greengard, Judith
;    APPLICANT:  Lee, Will
;    TITLE OF INVENTION:  METHODS FOR ADMINISTRATION OF
;    TITLE OF INVENTION:  RECOMBINANTGENE DELIVERY VEHICLES FOR TREATMENT
;    TITLE OF INVENTION:  OF HEMOPHILIA AND OTHER DISORDERS
;    NUMBER OF SEQUENCES:  84
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed Intellectual Property Law Group
;      STREET:  701 Fifth Avenue, Suite 6300
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/001,039B
;      FILING DATE:  13-JAN-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  1155.005 / 930049.441C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1457 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-001-039B-47

  Query Match           100.0%;  Score 1666;  DB 4;  Length 1457;
  Best Local Similarity  100.0%;  Pred. No. 1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1126 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1185

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1186 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1245

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1246 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1305

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1306 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1365

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1366 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1425

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1426 LLTRYLRIHPQSW 1438


RESULT 3
US-08-683-839B-3
; Sequence 3, Application US/08683839B
; Patent No. 5744326
;  GENERAL INFORMATION:
;    APPLICANT:  Ill, Charles . R. et al.
;    TITLE OF INVENTION:  Use of Viral Cis-Acting Post-Transcriptional
;    TITLE OF INVENTION:  Regulatory Sequences To Increase Expression of
;    TITLE OF INVENTION:  Intronless Genes Containing Near-Consensus Splice Sites
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/683,839B
;      FILING DATE:  11-MARCH-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Remillard, Jane E.
;      REGISTRATION NUMBER:  38,872
;      REFERENCE/DOCKET NUMBER:  TTI-138
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1471 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-683-839B-3

  Query Match           100.0%;  Score 1666;  DB 1;  Length 1471;
  Best Local Similarity  100.0%;  Pred. No. 1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1199

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1259

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1319

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1379

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1439

Qy        301 LLTRYLRIHPQSW 313
                 |||||||||||||
Db       1440 LLTRYLRIHPQSW 1452


RESULT 4
US-08-882-083-2
; Sequence 2, Application US/08882083
; Patent No. 5869292
;  GENERAL INFORMATION:
;    APPLICANT:  VOORBERG, Johannes J.
;    TITLE OF INVENTION:  HYBRID PROTEINS WITH MODIFIED ACTIVITY
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/882,083
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/558,107
;      FILING DATE:  13-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ISACSON, John P.
;      REGISTRATION NUMBER:  33,715
;      REFERENCE/DOCKET NUMBER:  30472/212
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1661 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-882-083-2

  Query Match             100.0%;  Score 1666;  DB 2;  Length 1661;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1389

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1449

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1509

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1569

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1629

Qy        301 LLTRYLRIHPQSW 313
                 |||||||||||||
Db       1630 LLTRYLRIHPQSW 1642


RESULT 5
US-08-558-107-2
; Sequence 2, Application US/08558107
; Patent No. 5910481
;  GENERAL INFORMATION:
;    APPLICANT:  VOORBERG, Johannes J.
;    TITLE OF INVENTION:  HYBRID PROTEINS WITH MODIFIED ACTIVITY
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/558,107
;      FILING DATE:  13-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ISACSON, John P.
;      REGISTRATION NUMBER:  33,715
;      REFERENCE/DOCKET NUMBER:  30472/212
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1661 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-558-107-2

  Query Match             100.0%;  Score 1666;  DB 2;  Length 1661;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1389

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1449

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1509

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1569

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1629

Qy        301 LLTRYLRIHPQSW 313
                 |||||||||||||
Db       1630 LLTRYLRIHPQSW 1642
```

```
Db       1630 LLTRYLRIHPQSW 1642


RESULT 6
US-09-243-539-2
; Sequence 2, Application US/09243539
; Patent No. 6130203
;  GENERAL INFORMATION:
;    APPLICANT:  VOORBERG, Johannes J.
;    TITLE OF INVENTION:  HYBRID PROTEINS WITH MODIFIED ACTIVITY
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/243,539
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/558,107
;      FILING DATE:  13-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ISACSON, John P.
;      REGISTRATION NUMBER:  33,715
;      REFERENCE/DOCKET NUMBER:  30472/212
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1661 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-243-539-2

  Query Match             100.0%;  Score 1666;  DB 3;  Length 1661;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1330 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1389

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1390 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1449

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1450 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1509

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1510 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1569

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1570 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1629

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1630 LLTRYLRIHPQSW 1642


RESULT 7
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    APPLICANT:  Runge, Marschall S.
;    TITLE OF INVENTION:  Hybrid Human/Porcine Factor VIII
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/864,004B
;      FILING DATE:  07 APRIL 1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU106
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6508
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapien
;      TISSUE TYPE:  Liver cDNA sequence
US-07-864-004B-4

  Query Match             100.0%;  Score 1666;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300
```

```
Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 8
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    APPLICANT:  Runge, Marschall S.
;    TITLE OF INVENTION:  Hybrid Human/Porcine Factor VIII
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/251,937A
;      FILING DATE:  31-MAY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/864,004
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pratt, John S.
;      REGISTRATION NUMBER:  29,476
;      REFERENCE/DOCKET NUMBER:  EMU106DIV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6367
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapien
;      TISSUE TYPE:  Liver cDNA sequence
US-08-251-937A-4

  Query Match             100.0%;  Score 1666;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 9
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    APPLICANT:  Runge, Marschall S.
;    TITLE OF INVENTION:  Hybrid Human/Animal Factor VIII
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,133A
;      FILING DATE:  March 11, 1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/864,004
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU/76677
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-572-6508
;      TELEFAX:  404-572-6555
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapien
;      TISSUE TYPE:  Liver cDNA sequence
US-08-212-133A-2

  Query Match             100.0%;  Score 1666;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120
```

```
Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 10
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
;  GENERAL INFORMATION:
;    APPLICANT:  Emory University
;    TITLE OF INVENTION:  Hybrid Human/Animal Factor VIII
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,503
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pratt, John S.
;      REGISTRATION NUMBER:  29,476
;      REFERENCE/DOCKET NUMBER:  EMU106CIP(3)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6500
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapien
;      TISSUE TYPE:  Liver cDNA sequence
US-08-474-503-2

  Query Match             100.0%;  Score 1666;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 11
US-08-670-707A-2
; Sequence 2, Application US/08670707A
; Patent No. 5859204
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    TITLE OF INVENTION:  Hybrid Human/Animal Factor VIII
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/670,707A
;      FILING DATE:  26-JUN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US94/13200
;      FILING DATE:  15-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,133
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/864,004
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Greenlee, Lorance L.
;      REGISTRATION NUMBER:  27,894
;      REFERENCE/DOCKET NUMBER:  75-95F
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303/499-8080
;      TELEFAX:  303/499-8089
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      TISSUE TYPE:  Liver
US-08-670-707A-2
```

```
  Query Match           100.0%;  Score 1666;  DB 2;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 12
US-09-037-601-2
; Sequence 2, Application US/09037601
; Patent No. 6180371
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    TITLE OF INVENTION:  Hybrid Human/Animal Factor VIII
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Greenlee, Winner and Sullivan, P.C.
;      STREET:  5370 Manhattan Circle Suite 201
;      CITY:  Boulder
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/037,601
;      FILING DATE:  26-JUN-1996
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US94/13200
;      FILING DATE:  15-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,133
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/864,004
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ferber, Donna M.
;      REGISTRATION NUMBER:  33,878
;      REFERENCE/DOCKET NUMBER:  75-95F
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303/499-8080
;      TELEFAX:  303/499-8089
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;      TISSUE TYPE:  Liver
US-09-037-601-2

  Query Match           100.0%;  Score 1666;  DB 3;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 13
US-09-324-867-3
; Sequence 3, Application US/09324867A
; Patent No. 6251632
; GENERAL INFORMATION:
;  APPLICANT: Lillicrap, David
;  APPLICANT: Cameron, Cherie
;  APPLICANT: No. 6251632ley, Colleen
;  APPLICANT: Horrocks, L. Suzanne Hoyle
;  APPLICANT: Hough, Christine
;  TITLE OF INVENTION: Canine Factor VIII Gene, Protein and Methods of Use
;  FILE REFERENCE: 1669.0010002/JAG/BJD
;  CURRENT APPLICATION NUMBER: US/09/324,867A
;  CURRENT FILING DATE: 1999-06-03
;  EARLIER APPLICATION NUMBER: 09/035,141
;  EARLIER FILING DATE: 1998-03-059
;  EARLIER APPLICATION NUMBER: 60/039,953
;  EARLIER FILING DATE: 1997-03-06
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-324-867-3

  Query Match           100.0%;  Score 1666;  DB 3;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313
```

```
RESULT 14
US-09-315-179-2
; Sequence 2, Application US/09315179
; Patent No. 6376463
; GENERAL INFORMATION:
;  APPLICANT: Lollar, John S
;  TITLE OF INVENTION: Modified Factor VIII
;  FILE REFERENCE: 75-95H
;  CURRENT APPLICATION NUMBER: US/09/315,179
;  CURRENT FILING DATE: 1999-05-20
;  EARLIER APPLICATION NUMBER: U.S. 09/037,601
;  EARLIER FILING DATE: 1998-03-10
;  EARLIER APPLICATION NUMBER: U.S. 08/670,707
;  EARLIER FILING DATE: 1996-06-26
;  EARLIER APPLICATION NUMBER: PCT/US97/11155
;  EARLIER FILING DATE: 1997-06-26
;  EARLIER APPLICATION NUMBER: PCT/US94/13200
;  EARLIER FILING DATE: 1994-11-15
;  EARLIER APPLICATION NUMBER: U.S. 08/212,133
;  EARLIER FILING DATE: 1994-03-11
;  EARLIER APPLICATION NUMBER: U.S. 07/864,004
;  EARLIER FILING DATE: 1992-04-07
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-315-179-2

  Query Match             100.0%;  Score 1666;  DB 3;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313
```

```
RESULT 15
US-09-523-656-2
; Sequence 2, Application US/09523656
; Patent No. 6458563
; GENERAL INFORMATION:
;  APPLICANT: Lollar S., John
;  TITLE OF INVENTION: MODIFIED FACTOR VIII
;  FILE REFERENCE: 75-95I
;  CURRENT APPLICATION NUMBER: US/09/523,656
;  CURRENT FILING DATE: 2000-03-10
;  EARLIER APPLICATION NUMBER: 09/037,601
;  EARLIER FILING DATE: 1998-03-10
;  EARLIER APPLICATION NUMBER: 08/670,707
;  EARLIER FILING DATE: 1996-06-26
;  NUMBER OF SEQ ID NOS: 38
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-523-656-2

  Query Match             100.0%;  Score 1666;  DB 4;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-167;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313
```

```
Search completed: April 20, 2005, 17:22:32
Job time : 31.2153 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:21:09 ; Search time 67.4035 Seconds
(without alignments)
1543.422 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332
Perfect score: 1666
Sequence: 1 LIGEHLHAGMSTLFLVYSNK.....VNSLDPPLLTRYLRIHPQSW 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1666	100.0	1438	13	US-10-006-091-1	Sequence 1, Appli
2	1666	100.0	1438	13	US-10-047-257-1	Sequence 1, Appli
3	1666	100.0	1438	14	US-10-225-900-1	Sequence 1, Appli
4	1666	100.0	1459	15	US-10-239-498A-13	Sequence 13, Appl
5	1666	100.0	1471	13	US-10-095-718-2	Sequence 2, Appli
6	1666	100.0	1471	15	US-10-681-970-2	Sequence 2, Appli
7	1666	100.0	2332	9	US-09-957-641-2	Sequence 2, Appli
8	1666	100.0	2332	14	US-10-187-319-2	Sequence 2, Appli
9	1666	100.0	2332	14	US-10-131-510A-2	Sequence 2, Appli
10	1666	100.0	2332	15	US-10-445-235-2	Sequence 2, Appli
11	1666	100.0	2332	15	US-10-360-101-229	Sequence 229, App
12	1666	100.0	2332	15	US-10-239-498A-2	Sequence 2, Appli
13	1666	100.0	2332	16	US-10-466-998A-1	Sequence 1, Appli
14	1666	100.0	2332	16	US-10-721-997A-34	Sequence 34, Appl
15	1666	100.0	2351	14	US-10-132-829-4	Sequence 4, Appli
16	1666	100.0	2351	14	US-10-172-712-27	Sequence 27, Appl
17	1666	100.0	2351	14	US-10-133-907-4	Sequence 4, Appli
18	1666	100.0	2351	15	US-10-411-037-30	Sequence 30, Appl
19	1666	100.0	2351	15	US-10-411-026-30	Sequence 30, Appl
20	1666	100.0	2351	15	US-10-410-962-30	Sequence 30, Appl
21	1666	100.0	2351	15	US-10-411-049-30	Sequence 30, Appl
22	1666	100.0	2351	16	US-10-410-930-30	Sequence 30, Appl
23	1666	100.0	2351	16	US-10-410-997-30	Sequence 30, Appl
24	1666	100.0	2351	16	US-10-411-012-30	Sequence 30, Appl
25	1666	100.0	2351	16	US-10-287-994-30	Sequence 30, Appl
26	1666	100.0	2351	16	US-10-410-913-30	Sequence 30, Appl
27	1666	100.0	2351	17	US-10-410-980-30	Sequence 30, Appl
28	1660	99.6	1459	15	US-10-239-498A-15	Sequence 15, Appl
29	1657	99.5	1459	15	US-10-239-498A-4	Sequence 4, Appli
30	1522	91.4	1431	13	US-10-095-718-4	Sequence 4, Appli
31	1522	91.4	1431	15	US-10-681-970-4	Sequence 4, Appli
32	1495	89.7	2319	14	US-10-187-319-6	GENERAL INFORMA
33	1495	89.7	2319	14	US-10-131-510A-6	Sequence 6, Appli
34	1484	89.1	2351	17	US-10-741-600-1034	Sequence 1034, Ap
35	1467	88.1	1443	14	US-10-187-319-39	Sequence 39, Appl
36	1467	88.1	1443	14	US-10-131-510A-39	Sequence 39, Appl
37	1467	88.1	2114	16	US-10-721-997A-33	Sequence 33, Appl
38	1467	88.1	2133	14	US-10-187-319-37	Sequence 37, Appl
39	1467	88.1	2133	14	US-10-131-510A-37	Sequence 37, Appl
40	1055	63.3	218	15	US-10-038-252-5	Sequence 5, Appli
41	1006	60.4	216	17	US-10-741-600-1033	Sequence 1033, Ap
42	748	44.9	160	15	US-10-298-796-9	Sequence 9, Appli
43	745.5	44.7	2196	15	US-10-360-101-259	Sequence 259, App
44	745.5	44.7	2224	14	US-10-115-563-14	Sequence 14, Appl
45	745.5	44.7	2224	14	US-10-172-712-31	Sequence 31, Appl

ALIGNMENTS

```
RESULT 1
US-10-006-091-1
; Sequence 1, Application US/10006091
; Publication No. US20020102730A1
; GENERAL INFORMATION:
;  APPLICANT: Cho, Myung-Sam
;  APPLICANT:  Chan, Sham-Yuen
;  APPLICANT:  Kelsey, William
;  APPLICANT:  Yee, Helena
;  TITLE OF INVENTION: Expression System for Factor VIII
;  FILE REFERENCE: MSB-7255.1
;  CURRENT APPLICATION NUMBER: US/10/006,091
;  CURRENT FILING DATE:  2001-12-06
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1438
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Derived from
;   OTHER INFORMATION: human factor VIII sequence
US-10-006-091-1

  Query Match             100.0%;  Score 1666;  DB 13;  Length 1438;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-162;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1166

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1226
```

```
Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1286

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1346

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1406

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1407 LLTRYLRIHPQSW 1419


RESULT 2
US-10-047-257-1
; Sequence 1, Application US/10047257
; Publication No. US20020115152A1
; GENERAL INFORMATION:
;  APPLICANT: Cho, Myung-Sam
;  APPLICANT:  Chan, Sham-Yuen
;  APPLICANT:  Kelsey, William
;  APPLICANT:  Yee, Helena
;  TITLE OF INVENTION: Expression System for Factor VIII
;  FILE REFERENCE: MSB-7255.2
;  CURRENT APPLICATION NUMBER: US/10/047,257
;  CURRENT FILING DATE:  2002-01-15
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1438
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Derived from
;   OTHER INFORMATION: human factor VIII sequence
US-10-047-257-1

  Query Match             100.0%;  Score 1666;  DB 13;  Length 1438;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-162;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1166

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1226

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1286

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1346

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1406

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1407 LLTRYLRIHPQSW 1419


RESULT 3
US-10-225-900-1
; Sequence 1, Application US/10225900
; Publication No. US20030077752A1
; GENERAL INFORMATION:
;  APPLICANT: Cho, Myung-Sam
;  APPLICANT:  Chan, Sham-Yuen
;  APPLICANT:  Kelsey, William
;  APPLICANT:  Yee, Helena
;  TITLE OF INVENTION: Expression System for Factor VIII
;  FILE REFERENCE: MSB-7255
;  CURRENT APPLICATION NUMBER: US/10/225,900
;  CURRENT FILING DATE:  2002-08-22
;  PRIOR APPLICATION NUMBER: US/09/209,916
;  PRIOR FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1438
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Derived from
;   OTHER INFORMATION: human factor VIII sequence
US-10-225-900-1

  Query Match             100.0%;  Score 1666;  DB 14;  Length 1438;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-162;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1107 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1166

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1167 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1226

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1227 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1286

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1287 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1346

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1347 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1406

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1407 LLTRYLRIHPQSW 1419


RESULT 4
US-10-239-498A-13
; Sequence 13, Application US/10239498A
; Publication No. US20040023333A1
; GENERAL INFORMATION:
;  APPLICANT: Hauser, Charlotte
;  APPLICANT:  Horster, Andrea
;  APPLICANT:  Schroder, Carola
;  APPLICANT:  Lehnerer, Michael
;  TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in
;  TITLE OF INVENTION:  Human Cell Lines
;  FILE REFERENCE: 80977.0001
;  CURRENT APPLICATION NUMBER: US/10/239,498A
;  CURRENT FILING DATE:  2003-07-08
;  PRIOR APPLICATION NUMBER: PCT/EP01/03220
;  PRIOR FILING DATE: 2001-03-21
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 1459
;   TYPE: PRT
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: pTGF8-2hyg-s
US-10-239-498A-13

  Query Match            100.0%;  Score 1666;  DB 15;  Length 1459;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-162;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1128 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1187

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1188 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1247

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1248 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1307

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1308 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1367

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1368 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1427

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1428 LLTRYLRIHPQSW 1440


RESULT 5
US-10-095-718-2
; Sequence 2, Application US/10095718
; Publication No. US20020131956A1
; GENERAL INFORMATION:
;  APPLICANT: Walsh, Christopher
;  APPLICANT:  Chao, Hengjun
;  APPLICANT:  Burstein, Haim
;  APPLICANT:  Lynch, Carmel
;  APPLICANT:  Stepan, Tony
;  APPLICANT:  Munson, Keith
;  TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 35052/204375
;  CURRENT APPLICATION NUMBER: US/10/095,718
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 09/689,430
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/158,780
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1471
;   TYPE: PRT
;   ORGANISM: Homo sapiens B-domain deleted factor VIII
;   FEATURE:
;   OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-095-718-2

  Query Match            100.0%;  Score 1666;  DB 13;  Length 1471;
  Best Local Similarity  100.0%;  Pred. No. 7.3e-162;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1199

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1259

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1319

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1379

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1439

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1440 LLTRYLRIHPQSW 1452


RESULT 6
US-10-681-970-2
; Sequence 2, Application US/10681970
; Publication No. US20040062752A1
; GENERAL INFORMATION:
;  APPLICANT: Walsh, Christopher
;  APPLICANT:  Chao, Hengjun
;  APPLICANT:  Burstein, Haim
;  APPLICANT:  Lynch, Carmel
;  APPLICANT:  Stepan, Tony
;  APPLICANT:  Munson, Keith
;  TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 35052/204375
;  CURRENT APPLICATION NUMBER: US/10/681,970
;  CURRENT FILING DATE:  2003-10-09
;  PRIOR APPLICATION NUMBER: US/09/689,430
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/158,780
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1471
;   TYPE: PRT
;   ORGANISM: Homo sapiens B-domain deleted factor VIII
;   FEATURE:
;   OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-681-970-2

  Query Match            100.0%;  Score 1666;  DB 15;  Length 1471;
  Best Local Similarity  100.0%;  Pred. No. 7.3e-162;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1140 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 1199

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1200 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 1259

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1260 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1319

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1320 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 1379

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 1439
```

```
Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       1440 LLTRYLRIHPQSW 1452


RESULT 7
US-09-957-641-2
; Sequence 2, Application US/09957641
; Publication No. US20020182670A1
; GENERAL INFORMATION:
;  APPLICANT: Emory University
;  TITLE OF INVENTION: MODIFIED FACTOR VIII
;  FILE REFERENCE: 75-00
;  CURRENT APPLICATION NUMBER: US/09/957,641
;  CURRENT FILING DATE:  2001-09-16
;  PRIOR APPLICATION NUMBER: US 60/234047
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: US 60/236460
;  PRIOR FILING DATE: 2000-09-29
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-957-641-2

  Query Match           100.0%;  Score 1666;  DB 9;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-161;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 8
US-10-187-319-2
; Sequence 2, Application US/10187319
; Publication No. US20030068785A1
;   GENERAL INFORMATION:
;        APPLICANT: Lollar, John S.
;        TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
;        NUMBER OF SEQUENCES: 40
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
;             STREET: 5370 Manhattan Circle Suite 201
;             CITY: Boulder
;             STATE: Colorado
;             COUNTRY: USA
;             ZIP: 80303
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/187,319
;             FILING DATE: 27-Aug-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 09/523,656
;             FILING DATE: 2000-03-10
;             APPLICATION NUMBER: US 09/037,601
;             FILING DATE: 1998-03-10
;             APPLICATION NUMBER: WO PCT/US97/11155
;             FILING DATE: 1997-06-26
;             APPLICATION NUMBER: US 08/670,707
;             FILING DATE: 1996-06-26
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Greenlee, Lorance L.
;             REGISTRATION NUMBER: 27,894
;             REFERENCE/DOCKET NUMBER: 75-95K
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 303/499-8080
;             TELEFAX: 303/499-8089
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2332 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: <Unknown>
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ANTI-SENSE: NO
;        FRAGMENT TYPE: N-terminal
;        ORIGINAL SOURCE:
;             ORGANISM: Homo sapiens
;             TISSUE TYPE: Liver
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-187-319-2

  Query Match           100.0%;  Score 1666;  DB 14;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-161;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 9
US-10-131-510A-2
; Sequence 2, Application US/10131510A
; Publication No. US20030166536A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Lollar, John S
;  TITLE OF INVENTION: Modified Factor VIII
;  FILE REFERENCE: 75-95J
;  CURRENT APPLICATION NUMBER: US/10/131,510A
;  CURRENT FILING DATE:  2002-07-10
;  PRIOR APPLICATION NUMBER: U.S. 09/315,179
;  PRIOR FILING DATE: 1999-05-20
;  PRIOR APPLICATION NUMBER: U.S. 09/037,601
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: U.S. 08/670,707
;  PRIOR FILING DATE: 1996-06-26
;  PRIOR APPLICATION NUMBER: PCT/US97/11155
;  PRIOR FILING DATE: 1997-06-26
;  PRIOR APPLICATION NUMBER: PCT/US94/13200
;  PRIOR FILING DATE: 1994-11-15
;  PRIOR APPLICATION NUMBER: U.S. 08/212,133
;  PRIOR FILING DATE: 1994-03-11
;  PRIOR APPLICATION NUMBER: U.S. 07/864,004
;  PRIOR FILING DATE: 1992-04-07
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-131-510A-2

  Query Match             100.0%;  Score 1666;  DB 14;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-161;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 10
US-10-445-235-2
; Sequence 2, Application US/10445235
; Publication No. US20040005670A1
; GENERAL INFORMATION:
;  APPLICANT: Katherine A. High
;  APPLICANT:  Rodney M. Camire
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  TREATMENT OF HEMOPHILIA A
;  FILE REFERENCE: CHOP0176
;  CURRENT APPLICATION NUMBER: US/10/445,235
;  CURRENT FILING DATE:  2003-05-22
;  PRIOR APPLICATION NUMBER: 60/382,486
;  PRIOR FILING DATE: 2002-05-22
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-10-445-235-2

  Query Match             100.0%;  Score 1666;  DB 15;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-161;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 11
US-10-360-101-229
; Sequence 229, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
;  APPLICANT: Moll, Gert  N.
;  APPLICANT:  Leenhouts, Cornelis J.
;  TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
;  FILE REFERENCE: 2183-5673
;  CURRENT APPLICATION NUMBER: US/10/360,101
;  CURRENT FILING DATE:  2003-02-07
;  PRIOR APPLICATION NUMBER: EP 02077060.8
;  PRIOR FILING DATE: 2002-05-24
;  NUMBER OF SEQ ID NOS: 309
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 229
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: sequence of factor VIII
US-10-360-101-229

  Query Match             100.0%;  Score 1666;  DB 15;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-161;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180
```

```
Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 12
US-10-239-498A-2
; Sequence 2, Application US/10239498A
; Publication No. US20040023333A1
; GENERAL INFORMATION:
;  APPLICANT: Hauser, Charlotte
;  APPLICANT:  Horster, Andrea
;  APPLICANT:  Schroder, Carola
;  APPLICANT:  Lehnerer, Michael
;  TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in
;  TITLE OF INVENTION:  Human Cell Lines
;  FILE REFERENCE: 80977.0001
;  CURRENT APPLICATION NUMBER: US/10/239,498A
;  CURRENT FILING DATE:  2003-07-08
;  PRIOR APPLICATION NUMBER: PCT/EP01/03220
;  PRIOR FILING DATE: 2001-03-21
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-239-498A-2

  Query Match           100.0%;  Score 1666;  DB 15;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-161;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 13
US-10-466-998A-1
; Sequence 1, Application US/10466998A
; Publication No. US20040126856A1
; GENERAL INFORMATION:
;  APPLICANT: BAJAJ, S. Paul
;  APPLICANT:  FAY, Philip J.
;  TITLE OF INVENTION: Factor IXa: Factor VIIIa Interaction and Methods
;  TITLE OF INVENTION:  Therefor
;  FILE REFERENCE: 66153-41436
;  CURRENT APPLICATION NUMBER: US/10/466,998A
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: PCT/US02/01724
;  PRIOR FILING DATE: 2002-01-23
;  PRIOR APPLICATION NUMBER: US 60/263,431
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-466-998A-1

  Query Match           100.0%;  Score 1666;  DB 16;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-161;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2301 LLTRYLRIHPQSW 2313


RESULT 14
US-10-721-997A-34
; Sequence 34, Application US/10721997A
; Publication No. US20040197875A1
; GENERAL INFORMATION:
;  APPLICANT: Aventis Behring GmbH
;  TITLE OF INVENTION: Modified cDNA Factor VIII and its Derivatives
;  FILE REFERENCE: 2002/M018-A66
;  CURRENT APPLICATION NUMBER: US/10/721,997A
;  CURRENT FILING DATE:  2003-11-26
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-721-997A-34

  Query Match           100.0%;  Score 1666;  DB 16;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-161;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2001 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2060
```

```
Qy          61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2061 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2120

Qy         121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2180

Qy         181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2240

Qy         241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2300

Qy         301 LLTRYLRIHPQSW 313
               |||||||||||||
Db        2301 LLTRYLRIHPQSW 2313


RESULT 15
US-10-132-829-4
; Sequence 4, Application US/10132829
; Publication No. US20030044982A1
; GENERAL INFORMATION:
;  APPLICANT: Chien, Kenneth  R
;  APPLICANT:  Hoshijima, Masahiko
;  TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VII
;  TITLE OF INVENTION:  with vesicle vector
;  FILE REFERENCE: 6627-PA1170
;  CURRENT APPLICATION NUMBER: US/10/132,829
;  CURRENT FILING DATE:  2002-04-25
;  PRIOR APPLICATION NUMBER: 60/286,314
;  PRIOR FILING DATE: 2001-04-25
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 2351
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-132-829-4

  Query Match            100.0%;  Score 1666;  DB 14;  Length 2351;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-161;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2020 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2079

Qy          61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2080 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2139

Qy         121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2140 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2199

Qy         181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2200 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2259

Qy         241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2260 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2319

Qy         301 LLTRYLRIHPQSW 313
               |||||||||||||
Db        2320 LLTRYLRIHPQSW 2332


Search completed: April 20, 2005, 17:34:49
```

Job time : 69.4035 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:13:08 ; Search time 21.6931 Seconds
(without alignments)
1388.270 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332
Perfect score: 1666
Sequence: 1 LIGEHLHAGMSTLFLVYSNK.....VNSLDPPLLTRYLRIHPQSW 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1666	100.0	2351	1	EZHU	coagulation factor
2	1495	89.7	2319	2	A47004	coagulation factor
3	1467	88.1	2133	2	T42763	coagulation factor
4	1006	60.4	216	2	A44258	factor VIII-associ
5	745.5	44.7	2224	1	KFHU5	coagulation factor
6	741.5	44.5	2183	2	T42764	coagulation factor
7	738.5	44.3	2211	1	KFBO5	coagulation factor
8	621.5	37.3	427	2	JC4915	ags protein precur
9	618	37.1	463	1	A36479	milk fat globule m
10	606.5	36.4	409	2	T11743	pP47 protein - pig
11	586.5	35.2	401	2	S65138	glycoprotein antig
12	586.5	35.2	427	2	S74211	PAS-6/7 protein pr
13	406	24.4	927	1	JQ0948	A5 antigen precurs
14	381.5	22.9	218	2	A47285	milk fat globule p
15	277.5	16.7	3133	2	S52093	hemocytin - silkwo
16	222	13.3	845	2	JC5256	adipocyte transcri
17	211	12.7	1283	2	T13799	neurexin IV - frui
18	202	12.1	737	2	T31349	hypothetical prote
19	191	11.5	737	2	T15615	hypothetical prote
20	185.5	11.1	855	2	S42621	protein-tyrosine k
21	182	10.9	719	2	S51739	transcription repr
22	173.5	10.4	819	2	I48859	tyro 10 receptor k
23	170.5	10.2	913	2	A48280	receptor tyrosine
24	169.5	10.2	876	2	A49508	protein-tyrosine k
25	167	10.0	910	2	A53137	tyrosine kinase re
26	158	9.5	1385	2	T14158	neurexin IV - mous
27	156	9.4	1381	2	T31083	paranodin - rat
28	119	7.1	791	2	T16031	hypothetical prote
29	113	6.8	149	1	DLDOID	discoidin I chain
30	105.5	6.3	253	1	DLDOIA	discoidin I chain
31	105	6.3	253	1	DLDOIC	discoidin I chain
32	98.5	5.9	149	2	B03382	discoidin I chain
33	98	5.9	874	4	GNHUER	retrovirus-related
34	94.5	5.7	159	2	AB0486	probable exported
35	91.5	5.5	452	2	A35721	nicotinic acetylch
36	91.5	5.5	1740	2	T43773	hypothetical prote
37	90.5	5.4	1827	2	T34288	hypothetical prote
38	90	5.4	531	1	SYNCYT	tyrosine-tRNA liga
39	89	5.3	815	2	F90951	biotin sulfoxide r
40	89	5.3	815	2	B85800	biotin sulfoxide r
41	88.5	5.3	883	1	JDVLC2	DNA-directed DNA p
42	88	5.3	1070	2	S46755	hypothetical prote
43	87	5.2	4377	2	A55575	ankyrin 3, long sp
44	86	5.2	399	1	JU0348	47.6K protein - Ch
45	86	5.2	455	2	E82658	beta-lactamase XF1

ALIGNMENTS

RESULT 1
EZHU
coagulation factor VIII precursor [validated] - human
N;Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant compor
C;Species: Homo sapiens (man)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004
C;Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S63527; S66445; B42
R;Gitschier, J.; Wood, W.I.
Hum. Mol. Genet. 1, 199-200, 1992
A;Title: Sequence of the exon-containing regions of the human factor VIII gene.
A;Reference number: I54318; MUID:93265012; PMID:1303178
A;Accession: I54318
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1921,'S',1923-2351 <RES>
A;Cross-references: UNIPROT:P00451; GB:M88648; NID:g182381; PIDN:AAA52420.1; PID:g182383
R;Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschier, J.; Keyt, B.; Seeburg,
Nature 312, 330-337, 1984
A;Title: Expression of active human factor VIII from recombinant DNA clones.
A;Reference number: A00525; MUID:85061548; PMID:6438526
A;Accession: A00525
A;Molecule type: mRNA
A;Residues: 1-2351 <WOO>
A;Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179
R;Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.D.; H
s, D.N.; Hewick, R.M.
Nature 312, 342-347, 1984
A;Title: Molecular cloning of a cDNA encoding human antihaemophilic factor.
A;Reference number: I58059; MUID:85061550; PMID:6438528
A;Accession: I58059
A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-74,'V',76-1259,'E',1261-2351 <RE2>
A;Cross-references: GB:K01740; NID:g182802; PIDN:AAA52484.1; PID:g182803
R;Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; Kuo,
.B.; Randolph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalaro, J.; Hansen, J.; No
DNA 4, 333-349, 1985
A;Title: Characterization of the polypeptide composition of human factor VIII:C and the
A;Reference number: A23584; MUID:86081164; PMID:3935400
A;Accession: A23584
A;Molecule type: mRNA
A;Residues: 1-2351 <TRU>
A;Cross-references: GB:M14113; NID:g182817; PIDN:AAA52485.1; PID:g182818
R;Eaton, D.; Rodriguez, H.; Vehar, G.A.
Biochemistry 25, 505-512, 1986
A;Title: Proteolytic processing of human factor VIII. Correlation of specific cleavages
ity.
A;Reference number: A26174; MUID:86159740; PMID:3082357
A;Accession: A26174
A;Molecule type: protein
A;Residues: 20-36;392-399,'X',401-402;1668-1678;1709-1722,'D',1723-1725;1741-1755 <EAT>
R;Pittman, D.D.; Wang, J.H.; Kaufman, R.J.

Biochemistry 31, 3315-3325, 1992
A;Title: Identification and functional importance of tyrosine sulfate residues within re
A;Reference number: A42348; MUID:92207952; PMID:1554716
A;Accession: A42348
A;Molecule type: protein
A;Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17
A;Experimental source: recombinant material from Chinese hamster ovary cells
A;Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
R;Fay, P.J.; Smudzin, T.M.
J. Biol. Chem. 264, 14005-14010, 1989
A;Title: Intersubunit fluorescence energy transfer in human factor VIII.
A;Reference number: A43986; MUID:89340500; PMID:2503509
A;Accession: A43986
A;Molecule type: protein
A;Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <FAY>
R;Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, K.;
J. Biol. Chem. 266, 740-746, 1991
A;Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
A;Reference number: A56109; MUID:91093266; PMID:1898735
A;Contents: annotation; sulfation
R;Gitschier, J.; Wood, W.I.; Goralka, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,
Nature 312, 326-330, 1984
A;Title: Characterization of the human factor VIII gene.
A;Reference number: A56196; MUID:85061547; PMID:6438525
A;Contents: annotation; introns
R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
Protein Sci. 4, 740-746, 1995
A;Title: Locations of disulfide bonds and free cysteines in the heavy and light chains c
A;Reference number: A56216; MUID:95338127; PMID:7613471
A;Contents: annotation; disulfide bonds
A;Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
R;Kjalke, M.; Heding, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
Eur. J. Biochem. 234, 773-779, 1995
A;Title: Amino acid residues 721-729 are required for full factor VIII activity.
A;Reference number: S63527; MUID:96163459; PMID:8575434
A;Accession: S63527
A;Molecule type: protein
A;Residues: 733-752;753-759 <KJA>
R;Lind, P.; Larsson, K.; Spira, J.; Sydow-Baeckman, M.; Almstedt, A.; Gray, E.; Sandberg
Eur. J. Biochem. 232, 19-27, 1995
A;Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
A;Reference number: S66445; MUID:96048024; PMID:7556150
A;Accession: S66445
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1668-1685 <LIN>
C;Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pr
C;Genetics:
A;Gene: GDB:F8C
A;Cross-references: GDB:119124; OMIM:306700
A;Map position: Xq28-Xq28
A;Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
C;Function:
A;Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa prc
A;Pathway: blood coagulation
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C;Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-2351/Product: coagulation factor VIII #status experimental <MAT>
F;20-740/Product: coagulation factor VIIIa heavy chain #status experimental <ACH>
F;20-356/Domain: A1 <DA1>
F;23-348/Domain: ferroxidase repeat homology <FO1>
F;392-759/Domain: A2 <DA2>
F;402-730/Domain: ferroxidase repeat homology <FO2>
F;760-1667/Domain: B <DB0>
F;1668-2351/Product: coagulation factor VIIIa light chain #status experimental <ACL>
F;1709-2038/Domain: A3 <DA3>
F;1716-2038/Domain: ferroxidase repeat homology <FO3>
F;2039-2191/Domain: C1 <DC1>
F;2039-2188/Domain: discoidin I amino-terminal homology <DN1>
F;2192-2351/Domain: C2 <DC2>
F;2192-2345/Domain: discoidin I amino-terminal homology <DN2>
F;60,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1
F;172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #status
F;355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted
F;365,737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experimental
F;391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F;414,426/Binding site: sulfate (Tyr) (covalent) #status predicted
F;759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F;1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
F;1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F;1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
F;2193-2345/Disulfide bonds: #status predicted

Query Match 100.0%; Score 1666; DB 1; Length 2351;
Best Local Similarity 100.0%; Pred. No. 2.1e-133;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2020 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2079

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2080 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2139

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2140 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2199

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2200 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2259

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2260 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2319

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2320 LLTRYLRIHPQSW 2332
```

RESULT 2
A47004
coagulation factor VIII precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: A47004
R;Elder, B.; Lakich, D.; Gitschier, J.
Genomics 16, 374-379, 1993
A;Title: Sequence of the murine factor VIII cDNA.
A;Reference number: A47004; MUID:93300511; PMID:8314577
A;Accession: A47004
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2319 <ELD>
A;Cross-references: UNIPROT:Q06194; GB:L05573; NID:g192456; PIDN:AAA37385.1; PID:g192457
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;23-349/Domain: ferroxidase repeat homology <FO1>
F;402-730/Domain: ferroxidase repeat homology <FO2>
F;1686-2006/Domain: ferroxidase repeat homology <FO3>
F;2007-2156/Domain: discoidin I amino-terminal homology <DN1>
F;2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 89.7%; Score 1495; DB 2; Length 2319;
Best Local Similarity 89.1%; Pred. No. 7.9e-119;
Matches 279; Conservative 13; Mismatches 21; Indels 0; Gaps 0;

```
Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||| |||||||||| :|| ||||||| |||||||||| |||||| ||||||||||
Db       1988 LIGEHLQAGMSTLFLVYSKQCQIPLGMASGSIRDFQITASGHYGQWAPNLARLHYSGSIN 2047

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||:|||||||||||||||||||||||||||||| :|:||||
```

```
Db       2048 AWSTKEPFSWIKVDLLAPMIVHGIKTQGARQKFSSLYISQFIIMYSLDGKKWLSYQGNST 2107

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||||||||||| |||||||||||||| |||||||||||||||||||:||||
Db       2108 GTLMVFFGNVDSSGIKHNSFNPPIIARYIRLHPTHSSIRSTLRMELMGCDLNSCSIPLGM 2167

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||| ||| |||||||||||||||||:|||||||:||||||||:||:||||| ||||||
Db       2168 ESKVISDTQITASSYFTNMFATWSPSQARLHLQGRTNAWRPQVNDPKQWLQVDLQKTMKV 2227

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||: ||||||| |||:||||||||||||| ||    |||||||||||| ||::||||||
Db       2228 TGIITQGVKSLFTSMFVKEFLISSSQDGHHWTQILYNGKVKVFQGNQDSSTPMMNSLDPP 2287

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||
Db       2288 LLTRYLRIHPQIW 2300


RESULT 3
T42763
coagulation factor VIII precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42763
R;Lollar, P.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z22269
A;Accession: T42763
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2133 <LOL>
A;Cross-references: UNIPROT:P12263; EMBL:U49517; NID:g1511633; PID:g1511634; PIDN:AAB067
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C;Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-2133/Product: coagulation factor VIII #status predicted <MAT>
F;23-349/Domain: ferroxidase repeat homology <FOX1>
F;402-730/Domain: ferroxidase repeat homology <FOX2>
F;1498-1820/Domain: ferroxidase repeat homology <FOX3>

  Query Match             88.1%;  Score 1467;  DB 2;  Length 2133;
  Best Local Similarity   86.3%;  Pred. No. 1.7e-116;
  Matches  270;  Conservative   25;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||| ||||| ||||| :|| ||||||| |||||||||||||||||||||||||||
Db       1802 LIGEHLQAGMSTTFLVYSKECQAPLGMASGRIRDFQITASGQYGQWAPKLARLHYSGSIN 1861

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||:| |||||||||||||||| |||||||||||||||||||: ||:||||||
Db       1862 AWSTKDPHSWIKVDLLAPMIIHGIMTQGARQKFSSLYISQFIIMYSLDGRNWQSYRGNST 1921

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||||:|||||||||||:|||||||||||||||||||||||||||||||||||
Db       1922 GTLMVFFGNVDASGIKHNIFNPPIVARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1981

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ::|||||:||||||: :|:|||||||:||||||||:|||||:|:: :|||||| |||:||
Db       1982 QNKAISDSQITASSHLSNIFATWSPSQARLHLQGRTNAWRPRVSSAEEWLQVDLQKTVKV 2041

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||:|||||||||:||||||||:|||||| :|||| |:|  |||||||| |||||:||||
Db       2042 TGITTQGVKSLLSSMYVKEFLVSSSQDGRRWTLFLQDGHTKVFQGNQDSSTPVVNALDPP 2101

Qy        301 LLTRYLRIHPQSW 313
              | |||||||| ||
Db       2102 LFTRYLRIHPTSW 2114


RESULT 4
A44258
factor VIII-associated gene B hypothetical protein - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: A44258
R;Levinson, B.; Kenwrick, S.; Gamel, P.; Fisher, K.; Gitschier, J.
Genomics 14, 585-589, 1992
A;Title: Evidence for a third transcript from the human factor VIII gene.
A;Reference number: A44258; MUID:93052386; PMID:1427887
A;Accession: A44258
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-216 <LEV>
A;Cross-references: UNIPROT:Q14286; GB:M90707; NID:g182316; PIDN:AAA58466.1; PID:g182317
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F;1-53/Domain: discoidin I amino-terminal homology #status atypical <DN1>
F;57-210/Domain: discoidin I amino-terminal homology <DN2>

  Query Match             60.4%;  Score 1006;  DB 2;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-78;
  Matches  189;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        125 VFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKA 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 VFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKA 68

Qy        185 ISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVT 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 ISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVT 128

Qy        245 TQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTR 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 TQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTR 188

Qy        305 YLRIHPQSW 313
              |||||||||
Db        189 YLRIHPQSW 197


RESULT 5
KFHU5
coagulation factor V precursor [validated] - human
N;Alternate names: coagulation labile factor; proaccelerin
C;Species: Homo sapiens (man)
C;Date: 19-May-1989 #sequence_revision 02-Jun-1995 #text_change 08-Dec-2000
C;Accession: A56172; A42344; A28028; A27498; A25897
R;Cripe, L.D.; Moore, K.D.; Kane, W.H.
Biochemistry 31, 3777-3785, 1992
A;Title: Structure of the gene for human coagulation factor V.
A;Reference number: A42344; MUID:92232668; PMID:1567832
A;Accession: A56172
A;Molecule type: DNA
A;Residues: 1-2224 <CRI>
A;Cross-references: GB:J05368
A;Accession: A42344
A;Molecule type: DNA
A;Residues: 48-58;79-89;120-130;191-201;239-249;313-323;368-378;428-437;461-471;533-542;5
2070;2111-2120;2172-2181 <CR2>
R;Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Kaufma
Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987
A;Title: Complete cDNA and derived amino acid sequence of human factor V.
A;Reference number: A28028; MUID:87260886; PMID:3110773
A;Accession: A28028
A;Molecule type: mRNA
A;Residues: 1-857,'R',859-864,'R',866-924,'E',926-1763,'I',1765-2212,'T',2214-2224 <JEN>
A;Cross-references: GB:M16967
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.
Biochemistry 26, 6508-6514, 1987
A;Title: Cloning of cDNAs coding for the heavy chain region and connecting region of huma
A;Reference number: A27498; MUID:88107560; PMID:2827731
A;Accession: A27498
A;Molecule type: mRNA
A;Residues: 1-1284,'I',1286-1600 <KAN>
A;Cross-references: GB:M17785
```

```
A;Note: parts of this sequence were determined by protein sequencing
R;Kane, W.H.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986
A;Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homolog
A;Reference number: A25897; MUID:86313665; PMID:3092220
A;Accession: A25897
A;Molecule type: mRNA
A;Residues: 1188-1215,1315-2224 <KA2>
A;Cross-references: GB:M14335
A;Note: parts of this sequence were determined by protein sequencing
R;Keller, F.G.; Ortel, T.L.; Quinn-Allen, M.A.; Kane, W.H.
Biochemistry 34, 4118-4124, 1995
A;Title: Thrombin-catalyzed activation of recombinant human factor V.
A;Reference number: A56139; MUID:95210278; PMID:7696276
A;Contents: annotation; thrombin cleavage sites
C;Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
C;Genetics:
A;Gene: GDB:F5
A;Cross-references: GDB:119896; OMIM:227400
A;Map position: 1q23-1q23
A;Introns: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1; 65
C;Function:
A;Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa prot
A;Pathway: blood coagulation
C;Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase re
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma;
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-2224/Product: coagulation factor V #status predicted <MAT>
F;29-737/Product: coagulation factor Va heavy chain #status experimental <VAH>
F;29-345/Domain: A1 <DA1>
F;33-329/Domain: ferroxidase repeat homology <FO1>
F;346-691/Domain: A2 <DA2>
F;351-684/Domain: ferroxidase repeat homology <FO2>
F;692-1573/Domain: B <DOB>
F;1183-1461/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)
F;1574-2224/Product: coagulation factor Va light chain #status experimental <VAL>
F;1574-1905/Domain: A3 <DA3>
F;1581-1905/Domain: ferroxidase repeat homology <FO3>
F;1667-1765/Region: phospholipid binding #status predicted
F;1906-2064/Domain: C1 <DC1>
F;1906-2061/Domain: discoidin I amino-terminal homology <DN1>
F;2065-2224/Domain: C2 <DC2>
F;2065-2221/Domain: discoidin I amino-terminal homology <DN2>
F;51,55,239,297,460,468,554,741,752,760,776,782,821,938,977,1074,1083,1103,1106,1479,149
F;167-193,248-329,500-526,603-684,1725-1751,1907-2061,2066-2221/Disulfide bonds: #status
F;334-335/Cleavage site: Arg-Asn (protein C) #status predicted
F;363,693,1546/Binding site: sulfate (Tyr) (covalent) #status predicted
F;376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F;382,1338/Binding site: carbohydrate (Asn) (covalent) #status absent
F;534-535/Cleavage site: Arg-Gly (protein C) #status predicted
F;737-738/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F;1046-1047/Cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experimenta
F;1573-1574/Cleavage site: Arg-Ser (thrombin) #status experimental

  Query Match             44.7%;  Score 745.5;  DB 1;  Length 2224;
  Best Local Similarity   44.9%;  Pred. No. 6.6e-55;
  Matches  144;  Conservative   65;  Mismatches  103;  Indels    9;  Gaps    4;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||:  ||| | ||:    |: |:|:::| | | || ||   | | |:||||:  || ||
Db       1888 VGENQRAGMQTPFLIMDRDCRMPMGLSTGIISDSQIKASEFLGYWEPRLARLNNGGSYNA 1947

Qy         62 WSTKE---PFS---WIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTY 115
              || ::    |:   ||:||:   :|| ||:||||:    | | ::| : || :   || :
Db       1948 WSVEKLAAEFASKPWIQVDMQKEVIITGIQTQGAKHYLKSCYTTEFYVAYSSNQINWQIF 2007

Qy        116 RGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCS 175
              :||||  :| | || |:| || | |:|||:|||||: ||    | |||:|| ||::| ||
Db       2008 KGNSTRNVMYFNGNSDASTIKENQFDPPIVARYIRISPTRAYNRPTLRLELQGCEVNGCS 2067

Qy        176 MPLGMESKAISDAQITASSYFTNMFAT-WSPSKARLHLQGRSNAWRPQVNNPKEWLQVDF 234
              |||||:  | : ||||||:  : :   | | :|||: ||| |||: : || |:||::|
Db       2068 TPLGMENGKIENKQITASSFKKSWWGDYWEPFRARLNAQGRVNAWQAKANNNKQWLEIDL 2127
```

```
Qy        235 QKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--KVFQGNQDSFTP 292
              |  |:| : ||| ||| : |||| : |  |: | :|  :     :  |:|:|| ::
Db       2128 LKIKKITAIITQGCKSLSSEMYVKSYTIHYSEQGVEWKPYRLKSSMVDKIFEGNTNTKGH 2187

Qy        293 VVNSLDPPLLTRYLRIHPQSW 313
              | |  :||:::|::|: |::|
Db       2188 VKNFFNPPIISRFIRVIPKTW 2208


RESULT 6
T42764
coagulation factor V - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42764
R;Yang, T.L.; Cui, J.; Rehumtulla, A.; Yang, A.; Moussalli, M.; Kaufman, R.J.; Ginsburg,
Blood 91, 4593-4599, 1998
A;Title: The structure and function of murine factor V and its inactivation by protein C.
A;Reference number: Z22270; MUID:98282202; PMID:9616155
A;Accession: T42764
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2183 <YAN>
A;Cross-references: UNIPROT:O88783; EMBL:U52925; NID:g3219690; PID:g3219691; PIDN:AAC995
C;Function:
A;Pathway: blood coagulation
C;Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase rep
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma; s
F;350-682/Domain: ferroxidase repeat homology <FOX1>
F;1541-1864/Domain: ferroxidase repeat homology <FOX2>

  Query Match             44.5%;  Score 741.5;  DB 2;  Length 2183;
  Best Local Similarity   45.2%;  Pred. No. 1.4e-54;
  Matches  145;  Conservative   60;  Mismatches  107;  Indels    9;  Gaps    4;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||:  ||| | ||:   :|: |:|:::| | | || ||      | |:||||: :|| ||
Db       1847 VGENQVAGMQTPFLIIDKECKMPMGLSTGVISDSQIKASEYLTYWEPRLARLNNAGSYNA 1906

Qy         62 WSTKE-----PFS-WIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTY 115
              || ::      |    ||:||:   ::: ||:||||:     | : ::| : || |   || :
Db       1907 WSIEKTALDFPIKPWIQVDMQKEVVVTGIQTQGAKHYLKSCFTTEFQVAYSSDQTNWQIF 1966

Qy        116 RGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCS 175
              || |  ::| | || | | || |  :|||:|||||:|||     | |||:|| ||::| ||
Db       1967 RGKSGKSVMYFTGNSDGSTIKENRLDPPIVARYIRIHPTKSYNRPTLRLELQGCEVNGCS 2026

Qy        176 MPLGMESKAISDAQITASSYFTNMFAT-WSPSKARLHLQGRSNAWRPQVNNPKEWLQVDF 234
              |||:|   | | ||||||:  : :   | || |||: ||| |||: : || |:|||||
Db       2027 TPLGLEDGRIQDKQITASSFKKSWWGDYWEPSLARLNAQGRVNAWQAKANNNKQWLQVDL 2086

Qy        235 QKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--KVFQGNQDSFTP 292
              |  ||| : ||| ||| : |||| : |  |  |  | :  |   :  |:|:|| ::
Db       2087 LKIKKVTAIVTQGCKSLSSEMYVKSYSIQYSDQGVAWKPYRQKSSMVDKIFEGNSNTKGH 2146

Qy        293 VVNSLDPPLLTRYLRIHPQSW 313
              : |  :||:::|::|| |::|
Db       2147 MKNFFNPPIISRFIRIIPKTW 2167


RESULT 7
KFBO5
coagulation factor V precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Mar-1993 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004
C;Accession: A42580; A36497
R;Guinto, E.R.; Esmon, C.T.; Mann, K.G.; MacGillivray, R.T.
J. Biol. Chem. 267, 2971-2978, 1992
A;Title: The complete cDNA sequence of bovine coagulation factor V.
A;Reference number: A42580; MUID:92147638; PMID:1737753
A;Accession: A42580
A;Molecule type: mRNA
```

```
A;Residues: 1-2211 <GUI>
A;Cross-references: UNIPROT:Q28107; GB:M81440; NID:g163037; PIDN:AAA30512.1; PID:g163038
A;Note: sequence extracted from NCBI backbone (NCBIN:80774, NCBIP:80776)
R;Kalafatis, M.; Jenny, R.J.; Mann, K.G.
J. Biol. Chem. 265, 21580-21589, 1990
A;Title: Identification and characterization of a phospholipid-binding site of bovine fa
A;Reference number: A36497; MUID:91072354; PMID:2254316
A;Accession: A36497
A;Molecule type: protein
A;Residues: 1566-1570,'X',1572-1581,'X',1583-1584;1673-1676,'X',1678-1679,'X',1681,'X',1
R;Xue, J.; Kalafatis, M.; Silveira, J.R.; Kung, C.; Mann, K.G.
Biochemistry 33, 13109-13116, 1994
A;Title: Determination of the disulfide bridges in factor Va heavy chain.
A;Reference number: A55979; MUID:95034740; PMID:7947716
A;Contents: annotation
A;Note: 566-Cys and 617-Cys were shown to have free sulfhydryls
C;Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
C;Function:
A;Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa prot
A;Pathway: blood coagulation
C;Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase re
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma;
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-2211/Product: coagulation factor V #status predicted <MAT>
F;29-741/Product: coagulation factor Va heavy chain #status predicted <VAH>
F;29-345/Domain: A1 <DA1>
F;33-329/Domain: ferroxidase repeat homology <FO1>
F;346-695/Domain: A2 <DA2>
F;351-688/Domain: ferroxidase repeat homology <FO2>
F;696-1564/Domain: B <DOB>
F;1175-1437/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)
F;1565-2211/Product: coagulation factor Va light chain #status predicted <VAL>
F;1565-1892/Domain: A3 <DA3>
F;1572-1892/Domain: ferroxidase repeat homology <FO3>
F;1654-1752/Region: phospholipid binding #status predicted
F;1893-2051/Domain: C1 <DC1>
F;1893-2048/Domain: discoidin I amino-terminal homology <DN1>
F;2052-2211/Domain: C2 <DC2>
F;2052-2208/Domain: discoidin I amino-terminal homology <DN2>
F;167-193,248-329,499-525/Disulfide bonds: #status experimental
F;225,239,297,382,460,553,587,745,756,774,780,902,952,964,1044,1053,1062,1071,1078,1094,
F;334-335/Cleavage site: Arg-Asn (protein C) #status predicted
F;363,697,1537/Binding site: sulfate (Tyr) (covalent) #status predicted
F;376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F;533-534/Cleavage site: Arg-Gly (protein C) #status predicted
F;607-688,1712-1738,1894-2048,2053-2208/Disulfide bonds: #status predicted
F;741-742/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predicted
F;1034-1035/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predicted
F;1564-1565/Cleavage site: Arg-Ser (thrombin) #status experimental

  Query Match             44.3%;  Score 738.5;  DB 1;  Length 2211;
  Best Local Similarity   44.5%;  Pred. No. 2.6e-54;
  Matches  143;  Conservative   61;  Mismatches  108;  Indels    9;  Gaps    4;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||   ||| | ||:   :|: |:|:::| | | || ||  :| | ||||||:  || ||
Db       1875 VGEIQRAGMQTPFLIVDRECKMPMGLSTGLIADSQIQASEFWGYWEPKLARLNNGGSYNA 1934

Qy         62 W-----STK-EPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTY 115
              |      ||:  |  ||:||:   ::: ||:||||:       | ::| : |||| | |: :
Db       1935 WIAEKLSTEFNPEPWIQVDMQKEVLLTGIQTQGAKHYLKPYYTTEFCVAYSLDRKNWRIF 1994

Qy        116 RGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCS 175
              :||||  :| | || |:| || |  :||::|||||: ||    :  ||:|| ||::| ||
Db       1995 KGNSTRNVMYFGGNSDASTIKENQIDPPVVARYIRISPTGSYNKPALRLELQGCEVNGCS 2054

Qy        176 MPLGMESKAISDAQITASSYFTNMFAT-WSPSKARLHLQGRSNAWRPQVNNPKEWLQVDF 234
              ||||||  | : |||||:  : :   | |  |||: ||| |||: : ||  :|||:|
Db       2055 TPLGMESGKIENKQITASSFKKSWWGNYWEPFLARLNAQGRVNAWQAKANNNNQWLQIDL 2114

Qy        235 QKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--KVFQGNQDSFTP 292
              |  |:| : ||| ||| : |||| : |  |  |  |  : :   :  |:|:|| :
Db       2115 LKIKKITAIVTQGCKSLSSEMYVKSYTIHYSDQGTDWKPYREKSSMVDKIFEGNNNVRGH 2174

Qy        293 VVNSLDPPLLTRYLRIHPQSW 313
              | |  :||:::|::|| |::|
Db       2175 VKNFFNPPIISRFIRIIPKTW 2195


RESULT 8
JC4915
ags protein precursor - rat
N;Alternate names: O-acetyl-Gd3 ganglioside
C;Species: Rattus norvegicus (Norway rat)
C;Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C;Accession: JC4915
R;Ogura, K.; Nara, K.; Watanabe, Y.; Kohno, K.; Tai, T.; Sanai, Y.
Biochem. Biophys. Res. Commun. 225, 932-938, 1996
A;Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.
A;Reference number: JC4915; MUID:96374422; PMID:8780713
A;Accession: JC4915
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-427 <OGU>
A;Cross-references: UNIPROT:P70490; DDBJ:D84068; NID:g1620006; PIDN:BAA12210.1; PID:g162(
A;Experimental source: CST cell
C;Comment: This protein is required for the O-acetylation of disialoganglioside sialic ac
C;Genetics:
A;Gene: ags
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolog
F;1-21/Domain: signal sequence #status predicted <SIG>
F;28-60/Domain: EGF homology <EG1>
F;68-107/Domain: EGF homology <EG2>
F;110-267/Domain: discoidin I amino-terminal homology <DN1>
F;271-427/Domain: discoidin I amino-terminal homology <DN2>

  Query Match             37.3%;  Score 621.5;  DB 2;  Length 427;
  Best Local Similarity   42.4%;  Pred. No. 2.9e-45;
  Matches  129;  Conservative   53;  Mismatches  111;  Indels   11;  Gaps    4;

Qy         21 CQTPLGMASGHIRDFQITASGQY------GQWAPKLARLHYSGSINAW--STKEPFSWIK 72
              | | ||:  | | | ||:||  |       :| |:||||: :| :|||  |: :   ||:
Db        111 CSTKLGLEGGAIADSQISASSVYMGFMGLQRWGPELARLYRTGIVNAWTASSYDSKPWIQ 170

Qy         73 VDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDS 132
              || |  | : |: |||| :   : |:  | : ||||:::  :  |       | || |:
Db        171 VDFLRKMRVSGVMTQGASRAGRAEYLKTFKVAYSLDGRRFEFIQDESGTGDKEFMGNQDN 230

Qy        133 SGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITA 192
              : :| |:||| : |:||||:|       ||| ||:||:|: || |||:::  | |:||||
Db        231 NSLKINMFNPTLEAQYIRLYPVSCHRGCTLRFELLGCELHGCSEPLGLKNNTIPDSQITA 290

Qy        193 SSYFT--NMFA-TWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVK 249
              || :   |: |  | |    ||  ||: |||  | |: |||||||     ||||: ||| :
Db        291 SSSYKTWNLRAFGWYPHLGRLDNQGKINAWTAQSNSAKEWLQVDLGTQKKVTGIITQGAR 350

Qy        250 SLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIH 309
                        ||  : :: | || |||:: : |  ||||||  |: :   |  : | : ||:|:
Db        351 DFGHIQYVASYKVAHSDDGVQWTVYEEQGTSKVFQGNLDNNSHKKNIFEKPFMARYVRVL 410

Qy        310 PQSW 313
              | ||
Db        411 PLSW 414


RESULT 9
A36479
milk fat globule membrane protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A36479
R;Stubbs, J.D.; Lekutis, C.; Singer, K.L.; Bui, A.; Yuzuki, D.; Srinivasan, U.; Parry, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990
A;Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the exi
A;Reference number: A36479; MUID:91046008; PMID:2122462
A;Accession: A36479
```

```
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-463 <STU>
A;Cross-references: UNIPROT:P21956; GB:M38337; NID:g199142; PIDN:AAA39534.1; PID:g199143
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
C;Keywords: membrane protein
F;28-60/Domain: EGF homology <EG1>
F;68-107/Domain: EGF homology <EG2>
F;147-303/Domain: discoidin I amino-terminal homology <DN1>
F;307-463/Domain: discoidin I amino-terminal homology <DN2>

  Query Match             37.1%;  Score 618;  DB 1;  Length 463;
  Best Local Similarity   41.7%;  Pred. No. 6.4e-45;
  Matches  128;  Conservative   59;  Mismatches  108;  Indels   12;  Gaps    5;

Qy         18 SNKCQTPLGMASGHIRDFQITASGQY------GQWAPKLARLHYSGSINAW--STKEPFS 69
              :::| | |||  | | | ||:||  |       :| |:||||: :| :|||  |  :
Db        145 ASRCSTQLGMEGGAIADSQISASYVYMGFMGLQRWGPELARLYRTGIVNAWHASNYDSKP 204

Qy         70 WIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGN 129
              ||:|:||  | : |: |||| :   : |:  | : |||||:|::  :  | |    | ||
Db        205 WIQVNLLRKMRVSGVMTQGASRAGRAEYLKTFKVAYSLDGRKFEFIQDESGGD-KEFLGN 263

Qy        130 VDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQ 189
              :|:: :| |:||| : |:||||:|       ||| ||:||:|: |  |||:::  | |:|
Db        264 LDNNSLKVNMFNPTLEAQYIRLYPVSCHRGCTLRFELLGCELHGCLEPLGLKNNTIPDSQ 323

Qy        190 ITASSYFT--NMFA-TWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQ 246
              ::||| :   |: |  | |    ||  ||: ||| | |: |||||||     :|||: ||
Db        324 MSASSSYKTWNLRAFGWYPHLGRLDNQGKINAWTAQSNSAKEWLQVDLGTQRQVTGIITQ 383

Qy        247 GVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYL 306
              | :      ||: : :: | || |||:: : |  |||||| |: :   |  : | : ||:
Db        384 GARDFGHIQYVESYKVAHSDDGVQWTVYEEQGSSKVFQGNLDNNSHKKNIFEKPFMARYV 443

Qy        307 RIHPQSW 313
              |: | ||
Db        444 RVLPVSW 450


RESULT 10
T11743
pP47 protein - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11743
R;Ensslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmidtke, J.; Matsuda, T.; Toepfe
Biol. Reprod. 58, 1057-1064, 1998
A;Title: Molecular cloning and characterization of P47, a novel boar sperm-associated zo
A;Reference number: Z17325; MUID:98206817; PMID:9546740
A;Accession: T11743
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-409 <ENS>
A;Cross-references: UNIPROT:P79385; EMBL:Y11683; NID:g2652927; PIDN:CAA72379.1; PID:g265
A;Experimental source: testis
C;Function:
A;Description: may be involved in membrane remodeling and/or function as a zona pellucid
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
F;6-40/Domain: EGF homology <EGF>

  Query Match             36.4%;  Score 606.5;  DB 2;  Length 409;
  Best Local Similarity   40.8%;  Pred. No. 5.2e-44;
  Matches  125;  Conservative   55;  Mismatches  113;  Indels   13;  Gaps    4;

Qy         21 CQTPLGMASGHIRDFQITASGQY------GQWAPKLARLHYSGSINAW--STKEPFSWIK 72
              |  |||| :| | ||||:||  :       :|||:|||| :| :|||  |  :   ||:
Db         91 CNAPLGMETGAIADFQISASSMHLGFMGLQRWAPELARLHRAGIVNAWTASNYDRNPWIQ 150

Qy         73 VDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDS 132
              |:||  | : |: |||| :  |: |:  | : || ||:|:|  :|      :| ||:|:
Db        151 VNLLRRMRVTGVVTQGASRAGSAEYMKTFKVAYSTDGRKFQFIQGAEESGDKIFMGNLDN 210

Qy        133 SGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITA 192
              ||:| |:|  |:  :|:|| |       ||| ||:||:|: |: |||::   | : ||||
Db        211 SGLKVNLFEVPLEVQYVRLVPIICHRGCTLRFELLGCELSGCAEPLGLKDNTIPNKQITA 270

Qy        193 SSYFTN---MFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVK 249
              ||::       :| |  |||  ||: |||  | |:  ||||:|     :|||: ||| :
Db        271 SSFYRTWGLSAFSWYPFYARLDNQGKFNAWTAQSNSASEWLQIDLGSQRRVTGIITQGAR 330

Qy        250 SLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--KVFQGNQDSFTPVVNSLDPPLLTRYLR 307
                    ||  : :: | ||  || :   | :  |:| || |: :   |  : | |||::|
Db        331 DFGHIQYVAAYKVAYSDDGVSWTEYRDQGALEGKIFPGNLDNNSHKKNMFETPFLTRFVR 390

Qy        308 IHPQSW 313
              | | :|
Db        391 ILPVAW 396


RESULT 11
S65138
glycoprotein antigen MGP57/53, mammary gland - bovine (fragment)
N;Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFG-E8 ho
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-Aug-1998
C;Accession: S65138; G48394
R;Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 385-391, 1995
A;Title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal ant
A;Reference number: S65138; MUID:96125736; PMID:8541316
A;Accession: S65138
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-401 <AOK>
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A;Reference number: A48394; MUID:93250576; PMID:8485470
A;Accession: G48394
A;Status: preliminary
A;Molecule type: protein
A;Residues: 207-220 <MAT>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131457)
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
C;Keywords: glycoprotein
F;1-32/Domain: EGF homology (fragment) <EG1>
F;40-79/Domain: EGF homology <EG2>
F;82-239/Domain: discoidin I amino-terminal homology <DN1>
F;243-401/Domain: discoidin I amino-terminal homology <DN2>

  Query Match             35.2%;  Score 586.5;  DB 2;  Length 401;
  Best Local Similarity   39.3%;  Pred. No. 2.6e-42;
  Matches  121;  Conservative   58;  Mismatches  112;  Indels   17;  Gaps    5;

Qy         21 CQTPLGMASGHIRDFQITASGQY------GQWAPKLARLHYSGSINAWST----KEPFSW 70
              | :|||| :| | | ||:||  :       :|||:|||| :| :|||::    | | |
Db         83 CTSPLGMQTGAIADSQISASSMHLGFMGLQRWAPELARLHQTGIVNAWTSGNYDKNP--W 140

Qy         71 IKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNV 130
              |:|:|:  | : |: |||| :  |: |:  | : || ||:::|  :         :| |||
Db        141 IQVNLMRKMWVTGVVTQGASRAGSAEYLKTFKVAYSTDGRQFQFIQVAGRSGDKIFIGNV 200

Qy        131 DSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQI 190
              ::||:| |:|: |:  :|:|| |       ||| ||:||:|| |: |||::   | : ||
Db        201 NNSGLKINLFDTPLETQYVRLVPIICHRGCTLRFELLGCELNGCTEPLGLKDNTIPNKQI 260

Qy        191 TASSYFTN---MFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQG 247
              |||||:        :| |  |||  ||: |||  | |:  ||||:|     :|||: |||
Db        261 TASSYYKTWGLSAFSWFPYYARLDNQGKFNAWTAQTNSASEWLQIDLGSQKRVTGIITQG 320

Qy        248 VKSLLTSMYVKEFLISSSQDGHQWTLFFQNG--KVKVFQGNQDSFTPVVNSLDPPLLTRY 305
              :       ||  : ::  ||  || :   | :  |:| || |: :   |  : | |   |:
Db        321 ARDFGHIQYVAAYRVAYGDDGVTWTEYKDPGASESKIFPGNMDNNSHKKNIFETPFQARF 380
```

```
Qy        306 LRIHPQSW 313
              :|| | :|
Db        381 VRIQPVAW 388


RESULT 12
S74211
PAS-6/7 protein precursor - bovine
N;Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFG-E8 h
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C;Accession: S74211; S78114; S24181; S65138; G48394
R;Hvarregaard, J.; Andersen, M.H.; Berglund, L.; Rasmussen, J.T.; Petersen, T.E.
Eur. J. Biochem. 240, 628-636, 1996
A;Title: Characterization of glycoprotein PAS-6/7 from membranes of bovine milk fat glob
A;Reference number: S74211; MUID:97008954; PMID:8856064
A;Accession: S74211
A;Molecule type: mRNA
A;Residues: 1-427 <HVA>
A;Cross-references: UNIPROT:Q95114; EMBL:X91895; NID:g1632778; PIDN:CAA62997.1; PID:g163
A;Accession: S78114
A;Molecule type: protein
A;Residues: 19-85;96-110;140-165;174-216;221-232;248-277;285-293;309-337;339-420;425-427
R;Kim, D.H.; Kanno, C.; Mizokami, Y.
Biochim. Biophys. Acta 1122, 203-211, 1992
A;Title: Purification and characterization of major glycoproteins, PAS-6 and PAS-7, from
A;Reference number: S23926; MUID:92353107; PMID:1643094
A;Accession: S24181
A;Molecule type: protein
A;Residues: 383-394 <KIM>
R;Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 385-391, 1995
A;Title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal an
A;Reference number: S65138; MUID:96125736; PMID:8541316
A;Accession: S65138
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 27-427 <AOK>
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A;Reference number: A48394; MUID:93250576; PMID:8485470
A;Accession: G48394
A;Status: preliminary
A;Molecule type: protein
A;Residues: 233-246 <MAT>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131457)
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
C;Keywords: blocked amino end; disulfide bond; glycoprotein; milk
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-427/Product: PAS-6/7 protein #status experimental <MAT>
F;24-58/Domain: EGF homology <EG1>
F;66-105/Domain: EGF homology <EG2>
F;108-265/Domain: discoidin I amino-terminal homology <DN1>
F;269-427/Domain: discoidin I amino-terminal homology <DN2>
F;24-35,29-47,49-58,66-77,71-94,96-105/Disulfide bonds: #status predicted
F;27/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;34/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;59,227/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;109-265,252-256,270-427/Disulfide bonds: #status experimental

  Query Match             35.2%;  Score 586.5;  DB 2;  Length 427;
  Best Local Similarity   39.3%;  Pred. No. 2.8e-42;
  Matches  121;  Conservative   58;  Mismatches  112;  Indels   17;  Gaps    5;

Qy         21 CQTPLGMASGHIRDFQITASGQY------GQWAPKLARLHYSGSINAWST----KEPFSW 70
              | :|||| :| | | ||:|| :       :|||:||||| :| :|||::    | |  |
Db        109 CTSPLGMQTGAIADSQISASSMHLGFMGLQRWAPELARLHQTGIVNAWTSGNYDKNP--W 166

Qy         71 IKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNV 130
              |:|:|:  | : |: |||| :  |: |:  | : || ||:::|  :        :| |||
Db        167 IQVNLMRKMWVTGVVTQGASRAGSAEYLKTFKVAYSTDGRQFQFIQVAGRSGDKIFIGNV 226

Qy        131 DSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQI 190
              ::||:| |:|: |:  :|:|| |       ||| ||:||:|| |: |||::   | : ||
Db        227 NNSGLKINLFDTPLETQYVRLVPIICHRGCTLRFELLGCELNGCTEPLGLKDNTIPNKQI 286

Qy        191 TASSYFTN---MFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQG 247
              |||||:        :| |  ||| ||: ||| | |:  ||||:|     :|||: |||
Db        287 TASSYYKTWGLSAFSWFPYYARLDNQGKFNAWTAQTNSASEWLQIDLGSQKRVTGIITQG 346

Qy        248 VKSLLTSMYVKEFLISSSQDGHQWTLFFQNG--KVKVFQGNQDSFTPVVNSLDPPLLTRY 305
               :      || : ::   ||  || :   |  : |:| || |: :   |  : |   |:
Db        347 ARDFGHIQYVAAYRVAYGDDGVTWTEYKDPGASESKIFPGNMDNNSHKKNIFETPFQARF 406

Qy        306 LRIHPQSW 313
              :|| | :|
Db        407 VRIQPVAW 414


RESULT 13
JQ0948
A5 antigen precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: JH0466; JQ0948
R;Takagi, S.; Hirata, T.; Agata, K.; Mochii, M.; Eguchi, G.; Fujisawa, H.
Neuron 7, 295-307, 1991
A;Title: The A5 antigen, a candidate for the neuronal recognition molecule, has homologie
A;Reference number: JH0466; MUID:91337458; PMID:1908252
A;Accession: JH0466
A;Molecule type: mRNA
A;Residues: 1-927 <TAK>
A;Cross-references: GB:D10467; GB:D01077; NID:g222962; PIDN:BAA01260.1; PID:g222963
A;Experimental source: tadpole, brain
A;Note: this protein has motifs homologous to complement components C1r and C1s and to cc
C;Comment: This protein is a neuronal cell surface molecule involved in the neuronal recc
C;Superfamily: Xenopus A5 antigen; C1r/C1s repeat homology; discoidin I amino-terminal hc
C;Keywords: duplication; glycoprotein; transmembrane protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-927/Product: A5 antigen #status predicted <A5A>
F;27-138/Domain: C1r/C1s repeat homology <C1R1>
F;147-262/Domain: C1r/C1s repeat homology <C1R2>
F;274-424/Domain: discoidin I amino-terminal homology <DN1>
F;430-584/Domain: discoidin I amino-terminal homology <DN2>
F;646-812/Domain: MAM homology <MAM>
F;861-883/Domain: transmembrane #status predicted <TMM>
F;150,261,300,523,844/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             24.4%;  Score 406;  DB 1;  Length 927;
  Best Local Similarity   34.4%;  Pred. No. 1.9e-26;
  Matches  108;  Conservative   54;  Mismatches  126;  Indels   26;  Gaps   12;

Qy         15 LVYSN-----KCQTPLGMASGHIRDFQITASGQYG-QWAPKLARLHYSGSINAWSTKEPF 68
              :| ||      :|:  ||| || |   ||: | ||   |: : :||:|    | |:  |
Db        264 VVQSNTDEDFQCKEALGMESGEIHFDQISVSSQYSMNWSAERSRLNYVE--NGWTPGEDT 321

Qy         69 --SWIKVDLLAPMIIHGIKTQGA--RQKFSSLYISQFIIMYSLDGKKWQTYR-GNSTGTL 123
                ||:|||     : || ||||  ::     ::  : :  | :|: | | : ||
Db        322 VKEWIQVDLENLRFVSGIGTQGAISKETKKKYFVKSYKVDISSNGEDWITLKDGNKH--- 378

Qy        124 MVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNS--CSMPLGME 181
              :|| || |:: : :  |: |:| |:|| |  :    :|| || || :    ||  |||
Db        379 LVFTGNTDATDVVYRPFSKPVITRFVRLRPVTWENGISLRFELYGCKITDYPCSRMLGMV 438

Qy        182 SKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNP--KEWLQVDFQKTMK 239
              |  |||:||||||    :   | |  |||        |  |   :|  |||||:|  :
Db        439 SGLISDSQITASS---QVDRNWVPELARLVTSRSGWALPPSNTHPYTKEWLQIDLAEEKI 495

Qy        240 VTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQN--GKVKVFQGNQDSFTPVVNSL 297
              | ||  || |     :::::| |  | :| :| :   :   | | |:|| :  || : :
Db        496 VRGVIIQGGKHKENKVFMRKFKIGYSNNGTEWEMIMDSSKNKPKTFEGNTNYDTPELRTF 555
```

```
Qy        298 DPPLLTRYLRIHPQ 311
                  : | ::|| |:
Db        556 -AHITTGFIRIIPE 568


RESULT 14
A47285
milk fat globule protein - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A47285
R;Larocca, D.; Peterson, J.A.; Urrea, R.; Kuniyoshi, J.; Bistrain, A.M.; Ceriani, R.L.
Cancer Res. 51, 4994-4998, 1991
A;Title: A M-r 46,000 human milk fat globule protein that is highly expressed in human b
A;Reference number: A47285; MUID:91371351; PMID:1909932
A;Accession: A47285
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-218 <LAR>
A;Cross-references: UNIPROT:Q08431; GB:S56151; NID:g235396; PIDN:AAB19771.1; PID:g235397
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
F;1-56/Domain: discoidin I amino-terminal homology (fragment) <DN1>
F;60-218/Domain: discoidin I amino-terminal homology <DN2>

  Query Match             22.9%;  Score 381.5;  DB 2;  Length 218;
  Best Local Similarity   39.2%;  Pred. No. 3.6e-25;
  Matches   76;  Conservative   40;  Mismatches   71;  Indels    7;  Gaps    3;

Qy        126 FFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKAI 185
              | || : : :  |:|  |: |:|:||:||      ||| ||:||:|| |: |||::: :|
Db         13 FVGNWNKNAVHVNLFETPVEAQYVRLYPTSCHTACTLRFELLGCELNGCANPLGLKNNSI 72

Qy        186 SDAQITASSYF----TNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVT 241
               | |||||| :     ::| :|:|| |||  ||  |||         :|||||   : :||
Db         73 PDKQITASSSYKTWGLHLF-SWNPSYARLDKQGNFNAWVAGSYGNDQWLQVDLGSSKEVT 131

Qy        242 GVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFF--QNGKVKVFQGNQDSFTPVVNSLDP 299
              |: ||| ::  :  :|  : :: | |   || :   : |  |:| || |: :   |  :
Db        132 GIITQGARNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFPGNWDNHSHKKNLFET 191

Qy        300 PLLTRYLRIHPQSW 313
              |:| ||:|| | :|
Db        192 PILARYVRILPVAW 205


RESULT 15
S52093
hemocytin - silkworm
N;Alternate names: humoral lectin
C;Species: Bombyx mori (silkworm)
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 12-Jul-2004
C;Accession: S52093; S70920
R;Kotani, E.; Yamakawa, M.; Iwamoto, S.; Tashiro, M.; Mori, H.; Sumida, M.; Matsubara, F
Biochim. Biophys. Acta 1260, 245-258, 1995
A;Title: Cloning and expression of the gene of hemocytin, an insect humoral lectin which
A;Reference number: S52093; MUID:95178544; PMID:7873598
A;Accession: S52093
A;Molecule type: mRNA
A;Residues: 1-3133 <KOT>
A;Cross-references: UNIPROT:P98092; EMBL:D29738
R;Mori, H.
submitted to the EMBL Data Library, April 1994
A;Reference number: S70920
A;Accession: S70920
A;Molecule type: mRNA
A;Residues: 1-1566,'S',1568-3133 <MOR>
A;Cross-references: EMBL:D29738; NID:g474967; PID:g664884
C;Superfamily: hemolectin/hemocytin

  Query Match             16.7%;  Score 277.5;  DB 2;  Length 3133;
  Best Local Similarity   25.9%;  Pred. No. 8.8e-15;
  Matches   82;  Conservative   61;  Mismatches  124;  Indels   49;  Gaps   13;

Qy         25 LGMASGHIRDFQITASGQYGQ-WAPKLARLH---YSGSINAWSTK--EPFSWIKVDLLAP 78
               | |     : |   :|| :: : :||  |||:     :   :|: |          :|:|:|
Db        948 LVMGDEPLPDTAFSASSEFSEIFAPHNARLNRGPTNSGAGSWNPKVNNDKQYIQVELPRR 1007

Qy         79 MIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHN 138
                |:|:  ||:        |:: : |||  ||  : |  |   |   :| | :|::
Db       1008 EPIYGVVLQGS--PIFDQYVTSYEIMYGDDGNTFSTVDG-PDGKPKIFRGPIDNTHPVKQ 1064

Qy        139 IFNPPIIARYIRLHPTHYSIRSTLRMELMGC---------------DLNSCSMPLGMES 182
              : :||| |: :|: |  :    :||:|::||                    |: |||:
Db       1065 MISPPIEAKVVRIRPLTWHDEISLRLEIIGCAEPLTTETSEPSPTSESPLQCTEPLGLIG 1124

Qy        183 K-AISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVT 241
              :  : : |::::          |  |  | : | :  |:|  | |  |:  ||     :|
Db       1125 ELPLENIQVSSN----------SEEKDYLSING-NRGWKPLYNTP-GWVMFDFTGPRNIT 1172

Qy        242 GVTTQGVK-SLLTS---MYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSL 297
              |: |:|     :||   :|  :|  :       :  ::|| |:|  | |    | |
Db       1173 GILTKGGNDGWVTSYKVLYTSDFETFN-------PVIDKDGKEKIFPANFDGIVSVTNEF 1225

Qy        298 DPPLLTRYLRIHPQSW 313
               ||:  |||:: || |
Db       1226 HPPIRARYLKVLPQKW 1241


Search completed: April 20, 2005, 17:21:42
Job time : 23.6931 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:11:58 ; Search time 78.25 Seconds
(without alignments)
2048.317 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2020_2332
Perfect score: 1666
Sequence: 1 LIGEHLHAGMSTLFLVYSNK.....VNSLDPPLLTRYLRIHPQSW 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1666	100.0	2351	1	FA8_HUMAN	P00451 homo sapien
2	1522	91.4	2343	2	O18806	O18806 canis famil
3	1517	91.1	2343	2	O62730	O62730 canis famil
4	1495	89.7	2319	1	FA8_MOUSE	Q06194 mus musculu
5	1467	88.1	2133	1	FA8_PIG	P12263 sus scrofa
6	1264	75.9	2258	2	Q7TN96	Q7tn96 rattus norv
7	1146	68.8	1377	2	Q804X3	Q804x3 gallus gall
8	1006	60.4	216	2	Q14286	Q14286 homo sapien
9	846	50.8	1639	2	Q804W6	Q804w6 fugu rubrip
10	782	46.9	745	2	Q804X4	Q804x4 gallus gall
11	771.5	46.3	1460	2	Q7SZN0	Q7szn0 pseudonaja
12	745.5	44.7	2224	1	FA5_HUMAN	P12259 homo sapien
13	745.5	44.7	2224	2	Q6UPU6	Q6upu6 homo sapien
14	743.5	44.6	2258	1	FA5_PIG	Q9glp1 sus scrofa
15	741.5	44.5	2183	2	O88783	O88783 mus musculu
16	738.5	44.3	2211	1	FA5_BOVIN	Q28107 bos taurus
17	721	43.3	407	2	Q8AYE0	Q8aye0 brachydanio
18	715.5	42.9	1802	2	Q804W5	Q804w5 fugu rubrip
19	685	41.1	2119	2	Q90X47	Q90x47 brachydanio
20	639.5	38.4	481	2	Q6RUW2	Q6ruw2 xenopus lae
21	625.5	37.5	480	1	EDI3_HUMAN	O43854 homo sapien
22	625.5	37.5	480	2	Q8N610	Q8n610 homo sapien
23	625	37.5	463	1	MFGM_MOUSE	P21956 mus musculu
24	621.5	37.3	427	1	MFGM_RAT	P70490 rattus norv
25	618.5	37.1	470	2	Q8C4U8	Q8c4u8 mus musculu
26	618.5	37.1	480	1	EDI3_MOUSE	O35474 mus musculu
27	618.5	37.1	480	2	Q8CBF7	Q8cbf7 mus musculu
28	606.5	36.4	409	1	MFGM_PIG	P79385 sus scrofa
29	586.5	35.2	427	1	MFGM_BOVIN	Q95114 bos taurus
30	584.5	35.1	475	2	Q6P1V9	Q6p1v9 xenopus tro
31	571.5	34.3	475	2	Q68EW5	Q68ew5 xenopus lae
32	544	32.7	387	1	MFGM_HUMAN	Q08431 homo sapien
33	520.5	31.2	2102	2	Q7TPK2	Q7tpk2 rattus norv
34	512	30.7	312	2	Q7Z3D2	Q7z3d2 homo sapien
35	506	30.4	363	2	O77718	O77718 equus cabal
36	466.5	28.0	384	2	Q8C8K0	Q8c8k0 mus musculu
37	444	26.7	335	2	Q9BTL9	Q9btl9 homo sapien
38	421	25.3	901	2	Q9H2D5	Q9h2d5 homo sapien
39	421	25.3	901	2	Q9H2E4	Q9h2e4 homo sapien
40	421	25.3	906	2	Q9H2D4	Q9h2d4 homo sapien
41	421	25.3	906	2	Q9H2E3	Q9h2e3 homo sapien
42	421	25.3	925	1	NRP2_RAT	O35276 rattus norv
43	421	25.3	931	1	NRP2_HUMAN	O60462 homo sapien
44	421	25.3	931	2	Q7Z3T9	Q7z3t9 homo sapien
45	419.5	25.2	926	2	Q8QZY7	Q8qzy7 mus musculu

ALIGNMENTS

RESULT 1
FA8_HUMAN
ID FA8_HUMAN STANDARD; PRT; 2351 AA.
AC P00451;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Coagulation factor VIII precursor (Procoagulant component)
DE (Antihemophilic factor) (AHF).
GN Name=F8; Synonyms=F8C;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86081164; PubMed=3935400;
RA Truett M.A., Blacher R., Burke R.L., Caput D., Chu C., Dina D.,
RA Hartog K., Kuo C.H., Masiarz F.R., Merryweather J.P., Najarian R.,
RA Pachl C., Potter S.J., Puma J., Quiroga M., Rall L.B., Randolph A.,
RA Urdea M.S., Valenzuela P., Dahl H.-H.M., Favalaro J., Hansen J.,
RA Nordfang O., Ezban M.;
RT "Characterization of the polypeptide composition of human factor
RT VIII:C and the nucleotide sequence and expression of the human kidney
RT cDNA.";
RL DNA 4:333-349(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061548; PubMed=6438526;
RA Wood W.I., Capon D.J., Simonsen C.C., Eaton D.L., Gitschier J.,
RA Keyt B., Seeburg P.H., Smith D.H., Hollingshead P., Wion K.L.,
RA Delwart E., Tuddenham E.G.D., Vehar G.A., Lawn R.M.;
RT "Expression of active human factor VIII from recombinant DNA clones.";
RL Nature 312:330-337(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061550; PubMed=6438528;
RA Toole J.J., Knopf J.L., Wozney J.M., Sultzman L.A., Buecker J.L.,
RA Pittman D.D., Kaufman R.J., Brown E., Shoemaker C., Orr E.C.,
RA Amphlett G.W., Foster W.B., Coe M.L., Knutson G.J., Fass D.N.,
RA Hewick R.M.;
RT "Molecular cloning of a cDNA encoding human antihaemophilic factor.";
RL Nature 312:342-347(1984).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93265012; PubMed=1303178;
RA Gitschier J., Wood W.I.;
RT "Sequence of the exon-containing regions of the human factor VIII
RT gene.";
RL Hum. Mol. Genet. 1:199-200(1992).
RN [5]
RP SEQUENCE OF 2064-2070 FROM N.A.
RA de Water N.S., Williams R., Browett P.J.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.

RN [6]
RP SULFATION OF TYR-1699.
RX MEDLINE=91093266; PubMed=1898735;
RA Leyte A., van Schijndel H.B., Niehrs C., Huttner W.B., Verbeet M.P.,
RA Mertens K., van Mourik J.A.;
RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
RT essential for the interaction of factor VIII with von Willebrand
RT factor.";
RL J. Biol. Chem. 266:740-746(1991).
RN [7]
RP SULFATION.
RX MEDLINE=92207952; PubMed=1554716;
RA Pittman D.D., Wang J.H., Kaufman R.J.;
RT "Identification and functional importance of tyrosine sulfate residues
RT within recombinant factor VIII.";
RL Biochemistry 31:3315-3325(1992).
RN [8]
RP SULFATION SITES TYR-737; TYR-738 AND TYR 742, AND DISULFIDE BONDS.
RX PubMed=7613471;
RA McMullen B.A., Fujikawa K., Davie E.W., Hedner U., Ezban M.;
RT "Locations of disulfide bonds and free cysteines in the heavy and
RT light chains of recombinant human factor VIII (antihemophilic factor
RT A).";
RL Protein Sci. 4:740-746(1995).
RN [9]
RP STRUCTURE BY NMR OF 2322-2343.
RX MEDLINE=95200924; PubMed=7893714;
RA Gilbert G.E., Baleja J.D.;
RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
RT amphipathic structure as determined by NMR spectroscopy.";
RL Biochemistry 34:3022-3031(1995).
RN [10]
RP REVIEW ON MOLECULAR BASIS OF HEMA.
RX MEDLINE=91221499; PubMed=1902642;
RA Gitschier J.;
RT "The molecular basis of hemophilia A.";
RL Ann. N. Y. Acad. Sci. 614:89-96(1991).
RN [11]
RP REVIEW ON MOLECULAR BASIS OF HEMA.
RX MEDLINE=89088506; PubMed=2491949;
RA White G.C. II, Shoemaker C.B.;
RT "Factor VIII gene and hemophilia A.";
RL Blood 73:1-12(1989).
RN [12]
RP REVIEW ON MOLECULAR BASIS OF HEMA.
RX MEDLINE=95245332; PubMed=7728145;
RA Antonarakis S.E., Kazazian H.H., Tuddenham E.G.D.;
RT "Molecular etiology of factor VIII deficiency in hemophilia A.";
RL Hum. Mutat. 5:1-22(1995).
RN [13]
RP VARIANT HEMA GLN-2326.
RX MEDLINE=86235434; PubMed=3012775;
RA Gitschier J., Wood W.I., Shuman M.A., Lawn R.M.;
RT "Identification of a missense mutation in the factor VIII gene of a
RT mild hemophiliac.";
RL Science 232:1415-1416(1986).
RN [14]
RP VARIANT HEMA PRO-2135.
RX MEDLINE=88096539; PubMed=3122181;
RA Levinson B., Janco R.L., Phillips J.A. III, Gitschier J.;
RT "A novel missense mutation in the factor VIII gene identified by
RT analysis of amplified hemophilia DNA sequences.";
RL Nucleic Acids Res. 15:9797-9805(1987).
RN [15]
RP VARIANT HEMA GLN-2228.
RX MEDLINE=88191889; PubMed=2833855;
RA Youssoufian H., Antonarakis S.E., Bell W., Griffin A.M.,
RA Kazazian H.H.;
RT "Nonsense and missense mutations in hemophilia A: estimate of the
RT relative mutation rate at CG dinucleotides.";
RL Am. J. Hum. Genet. 42:718-725(1988).
RN [16]
RP VARIANT HEMA GLY-291.
RX MEDLINE=88220354; PubMed=2835904;
RA Youssoufian H., Wong C., Aronis S., Platokoukis H., Kazazian H.H. Jr.,
RA Antonarakis S.E.;
RT "Moderately severe hemophilia A resulting from Glu-->Gly substitution
RT in exon 7 of the factor VIII gene.";
RL Am. J. Hum. Genet. 42:867-871(1988).
RN [17]
RP VARIANT HEMA CYS-1708.
RX MEDLINE=89274393; PubMed=2499363;
RA O'Brien D.P., Tuddenham E.G.;
RT "Purification and characterization of factor VIII 1,689-Cys: a
RT nonfunctional cofactor occurring in a patient with severe hemophilia
RT A.";
RL Blood 73:2117-2122(1989).
RN [18]
RP VARIANT HEMA CYS-391.
RX MEDLINE=90001543; PubMed=2506948;
RA Shima M., Ware J., Yoshioka A., Fukui H., Fulcher C.A.;
RT "An arginine to cysteine amino acid substitution at a critical
RT thrombin cleavage site in a dysfunctional factor VIII molecule.";
RL Blood 74:1612-1617(1989).
RN [19]
RP VARIANT HEMA LEU-189.
RX MEDLINE=90057680; PubMed=2510835;
RA Chan V., Chan T.K., Tong T.M., Todd D.;
RT "A novel missense mutation in exon 4 of the factor VIII:C gene
RT resulting in moderately severe hemophilia A.";
RL Blood 74:2688-2691(1989).
RN [20]
RP VARIANT HEMA LEU-2326.
RX MEDLINE=89197216; PubMed=2495245;
RA Inaba H., Fujimaki M., Kazazian H.H. Jr., Antonarakis S.E.;
RT "Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26
RT of the factor VIII gene.";
RL Hum. Genet. 81:335-338(1989).
RN [21]
RP VARIANT HEMA HIS-391.
RX MEDLINE=89264602; PubMed=2498882;
RA Arai M., Inaba H., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
RA Fujimaki M., Hoyer L.W.;
RT "Direct characterization of factor VIII in plasma: detection of a
RT mutation altering a thrombin cleavage site (arginine-
RT 372-->histidine).";
RL Proc. Natl. Acad. Sci. U.S.A. 86:4277-4281(1989).
RN [22]
RP VARIANT HEMA CYS-1708.
RX MEDLINE=90105723; PubMed=2104766;
RA Arai M., Higuchi M., Antonarakis S.E., Kazazian H.H. Jr.,
RA Phillips J.A. III, Janco R.L., Hoyer L.W.;
RT "Characterization of a thrombin cleavage site mutation (Arg 1689 to
RT Cys) in the factor VIII gene of two unrelated patients with cross-
RT reacting material-positive hemophilia A.";
RL Blood 75:384-389(1990).
RN [23]
RP VARIANTS HEMA GLN-2228 AND LEU-2326.
RX MEDLINE=90123183; PubMed=2105106;
RA Casula L., Murru S., Pecorara M., Ristaldi M.S., Restagno G.,
RA Mancuso G., Morfini M., de Biasi R., Baudo F., Carbonara A.;
RT "Recurrent mutations and three novel rearrangements in the factor VIII
RT gene of hemophilia A patients of Italian descent.";
RL Blood 75:662-670(1990).
RN [24]
RP VARIANT HEMA CYS-391.
RX MEDLINE=90329422; PubMed=1973901;
RA Pattinson J.K., McVey J.H., Boon M., Ajani A., Tuddenham E.G.;
RT "CRM+ haemophilia A due to a missense mutation (372-->Cys) at the
RT internal heavy chain thrombin cleavage site.";
RL Br. J. Haematol. 75:73-77(1990).
RN [25]
RP VARIANTS HEMA PHE-1699 AND CYS-1708.
RX MEDLINE=90152691; PubMed=2105906;
RA Higuchi M., Wong C., Kochhan L., Olek K., Aronis S., Kasper C.K.,
RA Kazazian H.H., Antonarakis S.E.;

```
RT   "Characterization of mutations in the factor VIII gene by direct
RT   sequencing of amplified genomic DNA.";

  Query Match             100.0%;  Score 1666;  DB 1;  Length 2351;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-132;
  Matches  313;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2020 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2079

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2080 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 2139

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2140 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 2199

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2200 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 2259

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2260 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 2319

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||||
Db       2320 LLTRYLRIHPQSW 2332


RESULT 2
O18806
ID   O18806      PRELIMINARY;      PRT;  2343 AA.
AC   O18806;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Factor VIII.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Cameron C., Notley C., Hoyle S., McGlynn L., Hough C., Kamisue S.,
RA   Giles A., Lillicrap D.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF016234; AAB87412.1; -.
DR   HSSP; P00451; 1D7P.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ   SEQUENCE   2343 AA;  265829 MW;  A854FAE571C3B399 CRC64;

  Query Match             91.4%;  Score 1522;  DB 2;  Length 2343;
  Best Local Similarity   89.8%;  Pred. No. 7.5e-120;
  Matches  281;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||||

Db       2012 LIGEHLQAGMSTLFLVYSKKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2071

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||:|||||||||||||||| ||||||||||:|||||||||| || :||||||
Db       2072 AWSTKDPFSWIKVDLLAPMIIHGIMTQGARQKFSSLYVSQFIIMYSLDGNKWHSYRGNST 2131

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||:|||||||||||||||||:||| |||||||||
Db       2132 GTLMVFFGNVDSSGIKHNIFNPPIIAQYIRLHPTHYSIRSTLRMELLGCDFNSCSMPLGM 2191

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              |||||||||||||| ::| |||||:||||||||:|||||| ||||||||||:|||||
Db       2192 ESKAISDAQITASSYLSSMLATWSPSQARLHLQGRTNAWRPQANNPKEWLQVDFRKTMKV 2251

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||:||||||||| |||||||||||||||| |||| ||||||||||:|| ||| | |:||
Db       2252 TGITTQGVKSLLISMYVKEFLISSSQDGHNWTLFLQNGKVKVFQGNRDSSTPVRNRLEPP 2311

Qy        301 LLTRYLRIHPQSW 313
              |: ||:|:|||||
Db       2312 LVARYVRLHPQSW 2324


RESULT 3
O62730
ID   O62730      PRELIMINARY;      PRT;  2343 AA.
AC   O62730;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Factor VIII.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Gordy P.W., Bowen R.A.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF049489; AAC05384.1; -.
DR   HSSP; P00451; 1IQD.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ   SEQUENCE   2343 AA;  265613 MW;  F612D744ADAADD99 CRC64;

  Query Match             91.1%;  Score 1517;  DB 2;  Length 2343;
  Best Local Similarity   89.5%;  Pred. No. 2e-119;
  Matches  280;  Conservative   17;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||| |||||||||||| ||||||||||||||||||||||||||||||||||||||||||

Db       2012 LIGEHLQAGMSTLFLVYSKKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 2071

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||:|||||||||||||||| ||||||||||:|||||||||| || :||||||
Db       2072 AWSTKDPFSWIKVDLLAPMIIHGIMTQGARQKFSSLYVSQFIIMYSLDGNKWHSYRGNST 2131

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              ||||||||||||||||||||||||:|||||||||||||||||:||| |||||||||
Db       2132 GTLMVFFGNVDSSGIKHNIFNPPIIAQYIRLHPTHYSIRSTLRMELLGCDFNSCSMPLGM 2191
```

```
Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||||||||||||||| ::| |||||||:||||||||:|||||| |||||||||||:|||||
Db       2192 ESKAISDAQITASSYLSSMLATWSPSQARLHLQGRTNAWRPQANNPKEWLQVDFRKTMKV 2251

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||:|||||||||| |||||||||||||||| |||| || ||||||||:|| ||| |:|:||
Db       2252 TGITTQGVKSLLISMYVKEFLISSSQDGHNWTLFLQNDKVKVFQGNRDSSTPVRNALEPP 2311

Qy        301 LLTRYLRIHPQSW 313
              |: ||:|:|||||
Db       2312 LVARYVRLHPQSW 2324


RESULT 4
FA8_MOUSE
ID   FA8_MOUSE      STANDARD;      PRT;  2319 AA.
AC   Q06194;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Coagulation factor VIII precursor (Procoagulant component).
GN   Name=F8; Synonyms=Cf8, F8c;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX   MEDLINE=93300511; PubMed=8314577;
RA   Elder B., Lakich D., Gitschier J.;
RT   "Sequence of the murine factor VIII cDNA.";
RL   Genomics 16:374-379(1993).
CC   -!- FUNCTION: Factor VIII, along with calcium and phospholipid, acts
CC       as a cofactor for factor IXa when it converts factor X to the
CC       activated form, factor Xa.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Found in most tissues.
CC   -!- PTM: The binding of Von Willibrand Factor (VWF) and activation
CC       depend on the sulfation of Tyr-1669.
CC   -!- SIMILARITY: Belongs to the multicopper oxidase family.
CC   -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L05573; AAA37385.1; -.
DR   PIR; A47004; A47004.
DR   HSSP; P00451; 1D7P.
DR   MGD; MGI:88383; F8.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW   Acute phase; Blood coagulation; Calcium; Glycoprotein; Plasma; Repeat;
KW   Signal; Sulfation.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20   2319       Coagulation factor VIII.
FT   DOMAIN       20    349       F5/8 type A 1.
FT   DOMAIN       20    199       Plastocyanin-like 1.
FT   DOMAIN      207    349       Plastocyanin-like 2.
FT   DOMAIN      399    730       F5/8 type A 2.
FT   DOMAIN      399    573       Plastocyanin-like 3.
FT   DOMAIN      583    730       Plastocyanin-like 4.
FT   DOMAIN      760   1640       B.
FT   DOMAIN     1683   2008       F5/8 type A 3.
FT   DOMAIN     1683   1845       Plastocyanin-like 5.
FT   DOMAIN     1855   2008       Plastocyanin-like 6.
FT   DOMAIN     2008   2156       F5/8 type C 1.
FT   DOMAIN     2161   2313       F5/8 type C 2.
FT   SITE        391    392       Cleavage (by thrombin) (By similarity).
FT   SITE        759    760       Cleavage (by thrombin) (By similarity).
FT   SITE       1324   1325       Cleavage (activation) (By similarity).
FT   SITE       1640   1641       Cleavage (activation) (By similarity).
FT   SITE       1678   1679       Cleavage (by thrombin) (By similarity).
FT   MOD_RES     367    367       Sulfotyrosine (By similarity).
FT   MOD_RES     737    737       Sulfotyrosine (By similarity).
FT   MOD_RES     738    738       Sulfotyrosine (By similarity).
FT   MOD_RES     742    742       Sulfotyrosine (By similarity).
FT   MOD_RES    1669   1669       Sulfotyrosine (By similarity).
FT   MOD_RES    1687   1687       Sulfotyrosine (By similarity).
FT   DISULFID    173    199       Probable.
FT   DISULFID    547    573       Probable.
FT   DISULFID   1819   1845       Probable.
FT   DISULFID   2008   2156       By similarity.
FT   DISULFID   2161   2313       By similarity.
FT   CARBOHYD     61     61       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    259    259       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    423    423       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    601    601       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    880    880       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    958    958       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1015   1015       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1022   1022       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1026   1026       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1044   1044       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1076   1076       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1087   1087       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1136   1136       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1161   1161       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1192   1192       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1255   1255       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1268   1268       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1273   1273       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1274   1274       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1302   1302       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1316   1316       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1340   1340       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1378   1378       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   1797   1797       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   2105   2105       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   2319 AA;  266148 MW;  FD054DE051DB2A01 CRC64;

  Query Match           89.7%;  Score 1495;  DB 1;  Length 2319;
  Best Local Similarity  89.1%;  Pred. No. 1.4e-117;
  Matches  279;  Conservative   13;  Mismatches   21;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||| |||||||||| :|| ||||||| |||||||||| |||||| ||||||||||
Db       1988 LIGEHLQAGMSTLFLVYSKQCQIPLGMASGSIRDFQITASGHYGQWAPNLARLHYSGSIN 2047

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||||||||||||||||||:|||||||||||||||||||||||||||||| :|:||||
Db       2048 AWSTKEPFSWIKVDLLAPMIVHGIKTQGARQKFSSLYISQFIIMYSLDGKKWLSYQGNST 2107

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||||||||||| |||||||||||||||| ||||||||||||||||||:||||
Db       2108 GTLMVFFGNVDSSGIKHNSFNPPIIARYIRLHPTHSSIRSTLRMELMGCDLNSCSIPLGM 2167

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||| ||| |||||||||||||||||:||||||||:||||||||:||:||||| ||||||
Db       2168 ESKVISDTQITASSYFTNMFATWSPSQARLHLQGRTNAWRPQVNDPKQWLQVDLQKTMKV 2227
```

```
Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||: ||||||| |||:|||||||||||||| ||    |||||||||||| ||::||||||
Db       2228 TGIITQGVKSLFTSMFVKEFLISSSQDGHHWTQILYNGKVKVFQGNQDSSTPMMNSLDPP 2287

Qy        301 LLTRYLRIHPQSW 313
              |||||||||||  |
Db       2288 LLTRYLRIHPQIW 2300


RESULT 5
FA8_PIG
ID   FA8_PIG        STANDARD;      PRT;  2133 AA.
AC   P12263; Q95243;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Coagulation factor VIII precursor (Procoagulant component).
GN   Name=F8; Synonyms=CF8;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Healey J.F., Lubin I.M., Lollar P.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 705-1573 FROM N.A.
RX   MEDLINE=86287369; PubMed=3016730;
RA   Toole J.J., Pittman D.D., Orr E.C., Murtha P., Wasley L.C.,
RA   Kaufman R.J.;
RT   "A large region (approximately equal to 95 kDa) of human factor VIII
RT   is dispensable for in vitro procoagulant activity.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5939-5942(1986).
RN   [3]
RP   SEQUENCE OF 392-759 FROM N.A.
RX   MEDLINE=94179260; PubMed=7510693;
RA   Lubin I.M., Healey J.F., Scandella D., Runge M.S., Lollar P.;
RT   "Elimination of a major inhibitor epitope in factor VIII.";
RL   J. Biol. Chem. 269:8639-8641(1994).
CC   -!- FUNCTION: Factor VIII, along with calcium and phospholipid, acts
CC       as a cofactor for factor IXa when it converts factor X to the
CC       activated form, factor Xa.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- SIMILARITY: Belongs to the multicopper oxidase family.
CC   -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U49517; AAB06705.1; -.
DR   PIR; A25945; A25945.
DR   PIR; T42763; T42763.
DR   HSSP; P00451; 1D7P.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
KW   Acute phase; Blood coagulation; Calcium; Glycoprotein; Plasma; Repeat;
KW   Signal; Sulfation.
FT   SIGNAL       1     19       Potential.
FT   CHAIN       20   2133       Coagulation factor VIII.
FT   DOMAIN      20    357       F5/8 type A 1.
FT   DOMAIN      20    199       Plastocyanin-like 1.
FT   DOMAIN     207    357       Plastocyanin-like 2.
FT   DOMAIN     399    730       F5/8 type A 2.
FT   DOMAIN     399    573       Plastocyanin-like 3.
FT   DOMAIN     583    730       Plastocyanin-like 4.
FT   DOMAIN     760   1599       B.
FT   DOMAIN    1495   1822       F5/8 type A 3.
FT   DOMAIN    1495   1659       Plastocyanin-like 5.
FT   DOMAIN    1669   1822       Plastocyanin-like 6.
FT   DOMAIN    1822   1970       F5/8 type C 1.
FT   DOMAIN    1975   2127       F5/8 type C 2.
FT   SITE       391    392       Cleavage (by thrombin) (By similarity).
FT   SITE       759    760       Cleavage (by thrombin) (By similarity).
FT   SITE      1449   1450       Cleavage (activation) (By similarity).
FT   SITE      1490   1491       Cleavage (by thrombin) (By similarity).
FT   MOD_RES    737    737       Sulfotyrosine (By similarity).
FT   MOD_RES    738    738       Sulfotyrosine (By similarity).
FT   MOD_RES    742    742       Sulfotyrosine (By similarity).
FT   DISULFID   173    199       Probable.
FT   DISULFID   547    573       Probable.
FT   DISULFID  1633   1659       Probable.
FT   DISULFID  1822   1970       By similarity.
FT   DISULFID  1975   2127       By similarity.
FT   CARBOHYD   233    233       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   259    259       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   601    601       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   929    929       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   985    985       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1025   1025       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1111   1111       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1181   1181       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1208   1208       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1245   1245       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1265   1265       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1335   1335       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1408   1408       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1611   1611       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  1919   1919       N-linked (GlcNAc. . .) (Potential).
FT   CONFLICT   713    713       N -> M (in Ref. 2).
FT   CONFLICT   734    734       I -> T (in Ref. 2).
FT   CONFLICT   792    792       G -> Q (in Ref. 2).
FT   CONFLICT  1133   1133       E -> F (in Ref. 2).
FT   CONFLICT  1191   1191       I -> L (in Ref. 2).
FT   CONFLICT  1209   1209       R -> F (in Ref. 2).
FT   CONFLICT  1437   1437       C -> G (in Ref. 2).
FT   CONFLICT  1456   1456       F -> R (in Ref. 2).
FT   CONFLICT  1539   1539       F -> R (in Ref. 2).
FT   CONFLICT  1546   1546       Q -> N (in Ref. 2).
SQ   SEQUENCE  2133 AA;  239304 MW;  152BBA8997F570DA CRC64;

  Query Match           88.1%;  Score 1467;  DB 1;  Length 2133;
  Best Local Similarity  86.3%;  Pred. No. 3.1e-115;
  Matches  270;  Conservative   25;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |||||| ||||| ||||| :|| ||||||| ||||||||||||||||||||||||||||
Db       1802 LIGEHLQAGMSTTFLVYSKECQAPLGMASGRIRDFQITASGQYGQWAPKLARLHYSGSIN 1861

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              |||||:| |||||||||||||||| ||||||||||||||||||||: ||:||||||
Db       1862 AWSTKDPHSWIKVDLLAPMIIHGIMTQGARQKFSSLYISQFIIMYSLDGRNWQSYRGNST 1921

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |||||||||:||||||||||||:||||||||||||||||||||||||||||||||||||
Db       1922 GTLMVFFGNVDASGIKHNIFNPPIVARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 1981

Qy        181 ESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ::|||||:|||||||: :|:||||||||:||||||||:|||||:|:|:: :|||||| |||:||
Db       1982 QNKAISDSQITASSHLSNIFATWSPSQARLHLQGRTNAWRPRVSSAEEWLQVDLQKTVKV 2041
```

```
Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||:|||||||||:|||||||:|||||| :|||| |:|  |||||||| |||||:||||
Db       2042 TGITTQGVKSLLSSMYVKEFLVSSSQDGRRWTLFLQDGHTKVFQGNQDSSTPVVNALDPP 2101

Qy        301 LLTRYLRIHPQSW 313
              | |||||||| ||
Db       2102 LFTRYLRIHPTSW 2114


RESULT 6
Q7TN96
ID   Q7TN96      PRELIMINARY;      PRT;  2258 AA.
AC   Q7TN96;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Factor VIII.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RA   Watzka M., Geisen C., Seifried E., Oldenburg J.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AY362193; AAQ21580.1; -.
DR   HSSP; P00451; 1CFG.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   Pfam; PF00394; Cu-oxidase; 1.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 3.
SQ   SEQUENCE   2258 AA;  251299 MW;  13AF91C788059B1D CRC64;

  Query Match           75.9%;  Score 1264;  DB 2;  Length 2258;
  Best Local Similarity   72.5%;  Pred. No. 5.6e-98;
  Matches  227;  Conservative   44;  Mismatches   42;  Indels    0;  Gaps    0;

Qy          1 LIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSIN 60
              |:|||| |||| |||||| :|| |||||||:||| |||||| :| | |:||||| :||:|
Db       1927 LVGEHLRAGMSALFLVYSTRCQVPLGMASGYIRDSQITASGYHGLWTPQLARLHNAGSVN 1986

Qy         61 AWSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNST 120
              ||| |||||:||||||||||:|||:||||| : ||||:|||||||||||:| :||||||
Db       1987 AWSAKEPFAWIKVDLLAPMILHGIETQGARHRLSSLYVSQFIIMYSLDGKQWLSYRGNST 2046

Qy        121 GTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGM 180
              |:|||||||||:| ::|| |:|||:|||||:|||| |||: |||||:|||||||| ||||
Db       2047 GSLMVFFGNVDASTVRHNRFDPPIVARYIRVHPTHASIRTALRMELLGCDLNSCSAPLGM 2106

Qy        181 ESKAISDAQITASSYFTNMPATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKV 240
              ||  ||| |:||||: :  |: | |: |||||:| :||||||||:| :||||| |:|:||
Db       2107 ESGWISDTQVTASSHLSTTFSAWPPALARLHLRGGANAWRPQVNDPTQWLQVDLQRTVKV 2166

Qy        241 TGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPP 300
              ||| ||| :||||:|:||:||:|:|||| ||   |:||||||||:|: ||:|||| ||
Db       2167 TGVVTQGARSLLTAMFVKKFLVSTSQDGRHWTHVLQDGKVKVFQGNRDASTPMVNSLHPP 2226

Qy        301 LLTRYLRIHPQSW 313
                |||||||| |
Db       2227 RFTRYLRIHPQVW 2239


RESULT 7
```

```
Q804X3
ID   Q804X3      PRELIMINARY;      PRT;  1377 AA.
AC   Q804X3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Coagulation factor VIII (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA   Tuddenham E.G.D., McVey J.H.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF465272; AAO33367.1; -.
DR   HSSP; P00451; 1D7P.
DR   GO; GO:0005507; F:copper ion binding; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 1.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   1377 AA;  152728 MW;  15ACFD7F765665CD CRC64;

  Query Match           68.8%;  Score 1146;  DB 2;  Length 1377;
  Best Local Similarity   67.0%;  Pred. No. 3.2e-88;
  Matches  209;  Conservative   46;  Mismatches   57;  Indels    0;  Gaps    0;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :|||| |||| |||||:  ||  |||||||:| | ||||||| | ||||||||| ||||||
Db       1047 VGEHLQAGMSALFLVYNQDCQNALGMASGYIADSQITASGQEGPWAPKLARLHNSGSINA 1106

Qy         62 WSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTG 121
              ||| |   ||:|||| ||||||||||||| ||||:|||| ||| |::|: |:||:|
Db       1107 WSTGESNPWIQVDLLHLKIIHGIKTQGARQKLSSLYVSQFIVFYSLHGQRWKKYKGNTTS 1166

Qy        122 TLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGME 181
              | |:|| |||:: :| | ||||||||||::||| |||:||||||:||||||||||||||
Db       1167 TQMMFFANVDATTVKENRFNPPIIARYIRINPTHDSIRTTLRMELIGCDLNSCSMPLGME 1226

Qy        182 SKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVT 241
              || | | :|:|||| :|:|::||||:|||:||||:||||: |:| |||||||: | |||
Db       1227 SKGIPDQRISASSYSSNIFSSWSPSQARLNLQGRTNAWRPETNSPSEWLQVDFEATKKVT 1286

Qy        242 GVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPL 301
              : ||| |:: | |:|||| :||:|||  |:    ||| |:|: |:|  : |:|||:|||
Db       1287 AIITQGAKAVFTHMFVKEFAVSSNQDGVHWSRVLHNGKEKIFRANRDYTSTVLNSLEPPL 1346

Qy        302 LTRYLRIHPQSW 313
                ||:||||: |
Db       1347 FARYVRIHPRHW 1358


RESULT 8
Q14286
ID   Q14286      PRELIMINARY;      PRT;  216 AA.
AC   Q14286;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Coagulation factor VIII associated protein B (Coagulation factor VIII,
DE   isoform b).
```

```
GN   Name=F8B;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93052386; PubMed=1427887;
RA   Levinson B., Kenwrick S., Gamel P., Fisher K., Gitschier J.;
RT   "Evidence for a third transcript from the human factor VIII gene.";
RL   Genomics 14:585-589(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; M90707; AAA58466.1; -.
DR   EMBL; BC022513; AAH22513.1; -.
DR   EMBL; BC064380; AAH64380.1; -.
DR   PIR; A44258; A44258.
DR   HSSP; P00451; 1CFG.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 1.
DR   PROSITE; PS01285; FA58C_1; 1.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
SQ   SEQUENCE   216 AA;  24641 MW;  6C82D4F89E35A376 CRC64;

  Query Match             60.4%;  Score 1006;  DB 2;  Length 216;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-77;
  Matches  189;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        125 VFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKA 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 VFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLGMESKA 68

Qy        185 ISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVT 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 ISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVT 128

Qy        245 TQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTR 304
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 TQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTR 188

Qy        305 YLRIHPQSW 313
              |||||||||
Db        189 YLRIHPQSW 197


RESULT 9
Q804W6
ID   Q804W6      PRELIMINARY;      PRT;  1639 AA.
AC   Q804W6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Coagulation factor VIII.
OS   Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Takifugu.
OX   NCBI_TaxID=31033;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA   Tuddenham E.G.D., McVey J.H.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF465279; AAO33374.1; -.
DR   HSSP; P00451; 1CFG.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_2.
SQ   SEQUENCE   1639 AA;  185203 MW;  881778AAFF809006 CRC64;

  Query Match             50.8%;  Score 846;  DB 2;  Length 1639;
  Best Local Similarity   50.0%;  Pred. No. 1.1e-62;
  Matches  160;  Conservative    47;  Mismatches  105;  Indels    8;  Gaps    2;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||   |||    |||: :|  |||| || | | || ||   | | | |||| || |||
Db       1301 VGESQLAGMRAKLLVYNPQCSRPLGMKSGRIGDSQIKASDYIGNWLPHLARLDQSGYINA 1360

Qy         62 WSTKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRG---- 117
              |      |||:||| || ::| ::||| |    : ||: | : |||| : | ||||
Db       1361 WMGYNEKSWIQVDLQAPTLLHRVQTQGVRSNLRNNYITAFTVSYSLDQETWSTYRGSGSS 1420

Qy        118 --NSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCS 175
                :|: |  || ||:|:| :|:| | || :|||||:|| :|: |  |||||:||||||||
Db       1421 RSSSSSTAKVFNGNLDNSRVKNNPFVPPFVARYIRIHPLYYNQRPALRMELLGCDLNSCS 1480

Qy        176 MPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQ 235
              :|||:: : | |    ||| : ::  :|:|| |||| :| :|||||: ||| ||||||
Db       1481 LPLGLQDRRIPDESFVASSSYWSLLRSWTPSLARLHQEGSANAWRPKNNNPHEWLQVDLG 1540

Qy        236 KTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNG--KVKVFQGNQDSFTPV 293
              |  ::||| ||| :|||| | | || :: |:||  |:   :    : |:|||| ||
Db       1541 KVKRITGVVTQGARSLLTKMMVTEFSVTISRDGQAWSSVLEGSSQREKIFQGNNDSDEEA 1600

Qy        294 VNSLDPPLLTRYLRIHPQSW 313
              :   | ||  ||:||||  |
Db       1601 LTIFDAPLFGRYIRIHPLGW 1620
```

```
RESULT 10
Q804X4
ID   Q804X4      PRELIMINARY;      PRT;   745 AA.
AC   Q804X4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Coagulation factor V (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Davidson C.J., Hirt R.P., Lal K., Snell P., Elgar G.,
RA   Tuddenham E.G.D., McVey J.H.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AF465271; AAO33366.1; -.
DR   HSSP; P12259; 1CZT.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   745 AA;  85626 MW;  A264587A348C29EE CRC64;

  Query Match             46.9%;  Score 782;  DB 2;  Length 745;
  Best Local Similarity   46.8%;  Pred. No. 1.2e-57;
  Matches  148;  Conservative   62;  Mismatches  102;  Indels    4;  Gaps    2;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||:  |||   :||    |: |:|||||| | | || ||     | ||||||: ||: ||
Db        412 VGEYQQAGMQASYLVLEKGCRIPMGMASGVILDSQINASHHIDYWEPKLARLNNSGTYNA 471

Qy         62 WST--KEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNS 119
              |||   |   ||:||    ::: ||:||||:    | |: :   |:|| | :|| |:||:|
Db        472 WSTTMTEDLPWIQVDFQRQVLLTGIQTQGAKHFLKSFYVQKLFIVYSTDKRKWNTFRGDS 531

Qy        120 TGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPLG 179
              :   :| || |: ||| || :|||||||||::|       ||||||:||::::||:|||
Db        532 SPAEKIFEGNSDAYGIKENIIDPPIIARYIRVYPKQAYNEPTLRMELLGCEVDACSLPLG 591

Qy        180 MESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMK 239
              ||:  | : ||||||  |: |:|| || ||:: :|: |||| : || ::|||:|     |
Db        592 MENGEIKNTQITASSVRTSWFSTWEPSLARINQKGKINAWRAESNNNQQWLQIDLLTIKK 651

Qy        240 VTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGK--VKVFQGNQDSFTPVVNSL 297
              :| : ||||||: |  :|| ::|  |  | :|  : ::     ||| ||:||   | :
Db        652 ITAIATQGVKSVTTENFVKTYVILYSNQGSEWKSYTEDSSSVAKVFSGNEDSRGHVKHFF 711

Qy        298 DPPLLTRYLRIHPQSW 313
              :||:|:|::|| |::|
Db        712 NPPILSRFIRIVPKTW 727


RESULT 11
Q7SZN0
ID   Q7SZN0      PRELIMINARY;      PRT;  1460 AA.
AC   Q7SZN0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Pseutarin C precursor.
OS   Pseudonaja textilis (Eastern brown snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Acanthophiinae; Pseudonaja.
OX   NCBI_TaxID=8673;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Venom gland;
RX   MEDLINE=22781770; PubMed=12730119; DOI=10.1182/blood-2002-12-3839;
RA   Rao V.S., Swarup S., Kini R.M.;
RT   "The nonenzymatic subunit of pseutarin C, a prothrombin activator from
RT   eastern brown snake (Pseudonaja textilis) venom, shows structural
RT   similarity to mammalian coagulation factor V.";
RL   Blood 102:1347-1354(2003).
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AY168281; AAO38805.1; -.
DR   HSSP; P12259; 1CZT.
DR   MEROPS; M14.951; -.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   InterPro; IPR000834; Peptidase_M14.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 1.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_3.
KW   Signal.
FT   SIGNAL        1     10       Potential.
SQ   SEQUENCE   1460 AA;  165931 MW;  6AFB63E2D5D275A6 CRC64;

  Query Match             46.3%;  Score 771.5;  DB 2;  Length 1460;
  Best Local Similarity   46.7%;  Pred. No. 2.1e-56;
  Matches  148;  Conservative   57;  Mismatches  107;  Indels    5;  Gaps    2;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||:   ||  || |    |: |:|:||| |:| ||:|||  | | ||||||: :|  ||
Db       1128 VGENQERGMQALFTVIDKDCKLPMGLASGIIQDSQISASGHVGYWEPKLARLNNTGKYNA 1187

Qy         62 WS---TKEPFSWIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGN 118
              ||    :    ||::||   ::| ||:||| |      |  :: : || ||: | |::|
Db       1188 WSIIKKEHEHPWIQIDLQRQVVITGIQTQGTVQLLQHSYTVEYFVTYSEDGQNWITFKGR 1247

Qy        119 STGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCSMPL 178
              : | | || | : :| |  :|||||||||||||| :  | | |:||:||::  ||:||
Db       1248 HSETQMHFEGNSDGTTVKENHIDPPIIARYIRLHPTKFYNRPTFRIELLGCEVEGCSVPL 1307

Qy        179 GMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTM 238
              |||| || :::|||||    :::| || |||:|:| :|||:|:|||  :|||:| |
Db       1308 GMESGAIKNSEITASSYKKTWWSSWEPSLARLNLEGGTNAWQPEVNNKDQWLQIDLQHLT 1367

Qy        239 KVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWT--LFFQNGKVKVFQGNQDSFTPVVNS 296
              |:| : |||  |: |||||| | |  : |   |   |  :    ||| || :|   | :
Db       1368 KITSIITQGATSMTTSMYVKTFSIHYTDDNSTWKPYLDVRTSMEKVFTGNINSDGHVKHF 1427

Qy        297 LDPPLLTRYLRIHPQSW 313
                ||:|:|::|| |::|
Db       1428 FKPPILSRFIRIIPKTW 1444


RESULT 12
FA5_HUMAN
ID   FA5_HUMAN      STANDARD;      PRT;  2224 AA.
AC   P12259; Q14285;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Coagulation factor V precursor (Activated protein C cofactor).
GN   Name=F5;
OS   Homo sapiens (Human).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANT MET-1764.
RX   MEDLINE=92232668; PubMed=1567832;
RA   Cripe L.D., Moore K.D., Kane W.H.;
RT   "Structure of the gene for human coagulation factor V.";
RL   Biochemistry 31:3777-3785(1992).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT GLU-925.
RX   MEDLINE=87260886; PubMed=3110773;
RA   Jenny R.J., Pittman D.D., Toole J.J., Kriz R.W., Aldape R.A.,
RA   Hewick R.M., Kaufman R.J., Mann K.G.;
RT   "Complete cDNA and derived amino acid sequence of human factor V.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:4846-4850(1987).
RN   [3]
RP   SEQUENCE OF 1-1600 FROM N.A., AND VARIANTS GLU-925 AND ILE-1285.
RX   MEDLINE=88107560; PubMed=2827731;
RA   Kane W.H., Ichinose A., Hagen F.S., Davie E.W.;
RT   "Cloning of cDNAs coding for the heavy chain region and connecting
RT   region of human factor V, a blood coagulation factor with four types
RT   of internal repeats.";
RL   Biochemistry 26:6508-6514(1987).
RN   [4]
RP   SEQUENCE OF 1188-1215 AND 1315-2224 FROM N.A.
RX   MEDLINE=86313665; PubMed=3092220;
RA   Kane W.H., Davie E.W.;
RT   "Cloning of a cDNA coding for human factor V, a blood coagulation
RT   factor homologous to factor VIII and ceruloplasmin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:6800-6804(1986).
RN   [5]
RP   PARTIAL SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=93203619; PubMed=8454869;
RA   Shen N.L.L., Fan S.-T., Pyati J., Graff R., Lapolla R.J.,
RA   Edgington T.S.;
RT   "The serine protease cofactor factor V is synthesized by
RT   lymphocytes.";
RL   J. Immunol. 150:2992-3001(1993).
RN   [6]
RP   SULFATION.
RX   MEDLINE=94264012; PubMed=8204629;
RA   Pittman D.D., Tomkinson K.N., Michnick D., Selighsohn U.,
RA   Kaufman R.J.;
RT   "Posttranslational sulfation of factor V is required for efficient
RT   thrombin cleavage and activation and for full procoagulant activity.";
RL   Biochemistry 33:6952-6959(1994).
RN   [7]
RP   SULFATION.
RX   MEDLINE=90366699; PubMed=2168225;
RA   Hortin G.L.;
RT   "Sulfation of tyrosine residues in coagulation factor V.";
RL   Blood 76:946-952(1990).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 2065-2224.
RX   MEDLINE=20052169; PubMed=10586886; DOI=10.1038/46594;
RA   Macedo-Ribeiro S., Bode W., Huber R., Quinn-Allen M.A., Kim S.W.,
RA   Ortel T.L., Bourenkov G.P., Bartunik H.D., Stubbs M.T., Kane W.H.,
RA   Fuentes-Prior P.;
RT   "Crystal structures of the membrane-binding C2 domain of human
RT   coagulation factor V.";
RL   Nature 402:434-439(1999).
RN   [9]
RP   VARIANT MET-1764.
RX   MEDLINE=95179146; PubMed=7874144;
RA   Bayston T.A., Ireland H., Olds R.J., Thein S.L., Lane D.A.;
RT   "A polymorphism in the human coagulation factor V gene.";
RL   Hum. Mol. Genet. 3:2085-2085(1994).
RN   [10]
RP   VARIANT APCR GLN-534.
RX   MEDLINE=94217810; PubMed=8164741; DOI=10.1038/369064a0;
RA   Bertina R.M., Koeleman B.P.C., Koster T., Rosendaal F.R., Dirven R.J.,
RA   de Ronde H., van der Velden P.A., Reitsma P.H.;
RT   "Mutation in blood coagulation factor V associated with resistance to
RT   activated protein C.";
RL   Nature 369:64-67(1994).
RN   [11]
RP   VARIANTS ILE-1285 AND ARG-1327.
RX   MEDLINE=96351768; PubMed=8713778;
RA   Lunghi B., Iacoviello L., Gemmati D., Dilasio M.G., Castoldi E.,
RA   Pinotti M., Castaman G., Redaelli R., Mariani G., Marchetti G.,
RA   Bernardi F.;
RT   "Detection of new polymorphic markers in the factor V gene:
RT   association with factor V levels in plasma.";
RL   Thromb. Haemost. 75:45-48(1996).
RN   [12]
RP   VARIANT APCR GLY-334, AND VARIANT LYS-513.
RX   MEDLINE=98122763; PubMed=9454741;
RA   Chan W.P., Lee C.K., Kwong Y.L., Lam C.K., Liang R.;
RT   "A novel mutation of Arg306 of factor V gene in Hong Kong Chinese.";
RL   Blood 91:1135-1139(1998).
RN   [13]
RP   VARIANT APCR THR-334.
RX   MEDLINE=98122764; PubMed=9454742;
RA   Williamson D., Brown K., Luddington R., Baglin C., Baglin T.;
RT   "Factor V Cambridge: a new mutation (Arg306-to-Thr) associated with
RT   resistance to activated protein C.";
RL   Blood 91:1140-1144(1998).
RN   [14]
RP   VARIANTS HIS-107; THR-413; LYS-513; SER-809; THR-817; ARG-858;
RP   ARG-865; GLU-925; GLN-1146; ALA-1530; SER-1685; VAL-1749; MET-1764;
RP   ILE-1820 AND GLY-2222, AND VARIANT APCR GLN-534.
RX   MEDLINE=99318093; PubMed=10391209; DOI=10.1038/10290;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RT   "Characterization of single-nucleotide polymorphisms in coding regions
RT   of human genes.";
RL   Nat. Genet. 22:231-238(1999).
RN   [15]
RP   ERRATUM.
RX   PubMed=10545957;
RA   Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA   Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA   Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA   Lander E.S.;
RL   Nat. Genet. 23:373-373(1999).
RN   [16]
RP   VARIANT APCR HIS-2102.
RX   MEDLINE=21847288; PubMed=11858490;
RA   Schrijver I., Houissa-Kastally R., Jones C.D., Garcia K.C.,
RA   Zehnder J.L.;
RT   "Novel factor V C2-domain mutation (R2074H) in two families with
RT   factor V deficiency and bleeding.";
RL   Thromb. Haemost. 87:294-299(2002).
CC   -!- FUNCTION: Coagulation factor V is a cofactor that participates
CC       with factor Xa to activate prothrombin to thrombin.
CC   -!- SUBUNIT: Factor Va is composed of a heavy chain and a light chain,
CC       noncovalently bound. The interaction between the two chains is
CC       calcium-dependent.
CC   -!- DOMAIN: Domain B contains 35 x 9 AA tandem repeats, and 2 x 17 AA
CC       repeats.
CC   -!- PTM: Thrombin activates factor V proteolytically to the active
CC       cofactor, factor Va (formation of a heavy chain at the N-terminus
CC       and a light chain at the C-terminus).
CC   -!- PTM: Sulfation is required for efficient thrombin cleavage and
CC       activation and for full procoagulant activity.
CC   -!- DISEASE: Defects in F5 are the cause of Owren parahemophilia
CC       [MIM:227400], an hemorrhagic diastesis.
CC   -!- DISEASE: Defects in F5 are the cause of resistance to activated
CC       protein C (APCR) [MIM:188055], a form of thrombophilia. The APCR
CC       mutation is found in about 5% of the population which suggest that
CC       a slight thrombotic tendency may confer some advantage in fetal
CC       implantation.
```

```
CC   -!- SIMILARITY: Belongs to the multicopper oxidase family.
CC   -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L32779; AAB59401.1; -.
DR   EMBL; L32755; AAB59401.1; JOINED.
DR   EMBL; L32756; AAB59401.1; JOINED.
DR   EMBL; L32757; AAB59401.1; JOINED.
DR   EMBL; L32758; AAB59401.1; JOINED.
DR   EMBL; L32759; AAB59401.1; JOINED.
DR   EMBL; L32760; AAB59401.1; JOINED.
DR   EMBL; L32761; AAB59401.1; JOINED.
DR   EMBL; L32762; AAB59401.1; JOINED.
DR   EMBL; L32763; AAB59401.1; JOINED.
DR   EMBL; L32764; AAB59401.1; JOINED.
DR   EMBL; L32765; AAB59401.1; JOINED.
DR   EMBL; L32766; AAB59401.1; JOINED.
DR   EMBL; L32767; AAB59401.1; JOINED.
DR   EMBL; L32768; AAB59401.1; JOINED.
DR   EMBL; L32769; AAB59401.1; JOINED.
DR   EMBL; L32770; AAB59401.1; JOINED.
DR   EMBL; L32771; AAB59401.1; JOINED.
DR   EMBL; L32772; AAB59401.1; JOINED.
DR   EMBL; L32773; AAB59401.1; JOINED.
DR   EMBL; L32774; AAB59401.1; JOINED.
DR   EMBL; L32775; AAB59401.1; JOINED.
DR   EMBL; L32776; AAB59401.1; JOINED.
DR   EMBL; L32777; AAB59401.1; JOINED.
DR   EMBL; L32778; AAB59401.1; JOINED.
DR   EMBL; M16967; AAA52424.1; -.
DR   EMBL; M14335; AAB59532.1; -.
DR   PDB; 1CZS; X-ray; A=2065-2224.
DR   PDB; 1CZT; X-ray; A=2065-2224.
DR   PDB; 1CZV; X-ray; A/B=2065-2224.
DR   PDB; 1FV4; Model; H=29-691, L=1574-2224.
DR   Genew; HGNC:3542; F5.
DR   MIM; 227400; -.
DR   MIM; 134400; -.

  Query Match             44.7%;  Score 745.5;  DB 1;  Length 2224;
  Best Local Similarity   44.9%;  Pred. No. 5.6e-54;
  Matches  144;  Conservative   65;  Mismatches  103;  Indels    9;  Gaps    4;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||:  ||| | ||:    |: |:|:::| | | || ||   | | |:||||:  || ||
Db       1888 VGENQRAGMQTPFLIMDRDCRMPMGLSTGIISDSQIKASEFLGYWEPRLARLNNGGSYNA 1947

Qy         62 WSTKE---PFS---WIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTY 115
              || ::    |:   ||:||:   :|| ||:||||:    | | ::| : || :   || :
Db       1948 WSVEKLAAEFASKPWIQVDMQKEVIITGIQTQGAKHYLKSCYTTEFYVAYSSNQINWQIF 2007

Qy        116 RGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCS 175
              :||||  :| | || |:| || | |:|||:|||||: ||    | |||:|| ||::| ||
Db       2008 KGNSTRNVMYFNGNSDASTIKENQFDPPIVARYIRISPTRAYNRPTLRLELQGCEVNGCS 2067

Qy        176 MPLGMESKAISDAQITASSYFTNMFAT-WSPSKARLHLQGRSNAWRPQVNNPKEWLQVDF 234
              |||||:  | : |||||:  : :   | | :|||: ||| |||: : || |:||::|
Db       2068 TPLGMENGKIENKQITASSFKKSWWGDYWEPFRARLNAQGRVNAWQAKANNNKQWLEIDL 2127

Qy        235 QKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--KVFQGNQDSFTP 292
              |  |:| : ||| ||| : |||| : |  |: | :|  :    :  |:|:|| ::
Db       2128 LKIKKITAIITQGCKSLSSEMYVKSYTIHYSEQGVEWKPYRLKSSMVDKIFEGNTNTKGH 2187

Qy        293 VVNSLDPPLLTRYLRIHPQSW 313
              | |  :||:::|::|: |::|
Db       2188 VKNFFNPPIISRFIRVIPKTW 2208


RESULT 13
Q6UPU6
ID   Q6UPU6       PRELIMINARY;      PRT;  2224 AA.
AC   Q6UPU6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Coagulation factor V (Proaccelerin, labile factor).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,
RA   Ahearn M.O., Kuldanek S.A., Rajkumar N., Toth E.J., Yi Q.,
RA   Nickerson D.A.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; AY364535; AAQ55063.1; -.
DR   HSSP; P00450; 1KCW.
DR   GO; GO:0007596; P:blood coagulation; IEA.
DR   GO; GO:0007155; P:cell adhesion; IEA.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR009271; LSPD.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   Pfam; PF06049; LSPR; 28.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_2.
SQ   SEQUENCE   2224 AA;  251701 MW;  24AEF3FEA7332E37 CRC64;

  Query Match             44.7%;  Score 745.5;  DB 2;  Length 2224;
  Best Local Similarity   44.9%;  Pred. No. 5.6e-54;
  Matches  144;  Conservative   65;  Mismatches  103;  Indels    9;  Gaps    4;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||:  ||| | ||:    |: |:|:::| | | || ||   | | |:||||:  || ||
Db       1888 VGENQRAGMQTPFLIMDRDCRMPMGLSTGIISDSQIKASEFLGYWEPRLARLNNGGSYNA 1947

Qy         62 WSTKE---PFS---WIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTY 115
              || ::    |:   ||:||:   :|| ||:||||:    | | ::| : || :   || :
Db       1948 WSVEKLAAEFASKPWIQVDMQKEVIITGIQTQGAKHYLKSCYTTEFYVAYSSNQINWQIF 2007

Qy        116 RGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCS 175
              :||||  :| | || |:| || | |:|||:|||||: ||    | |||:|| ||::| ||
Db       2008 KGNSTRNVMYFNGNSDASTIKENQFDPPIVARYIRISPTRAYNRPTLRLELQGCEVNGCS 2067

Qy        176 MPLGMESKAISDAQITASSYFTNMFAT-WSPSKARLHLQGRSNAWRPQVNNPKEWLQVDF 234
              |||||:  | : |||||:  : :   | | :|||: ||| |||: : || |:||::|
Db       2068 TPLGMENGKIENKQITASSFKKSWWGDYWEPFRARLNAQGRVNAWQAKANNNKQWLEIDL 2127

Qy        235 QKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--KVFQGNQDSFTP 292
              |  |:| : ||| ||| : |||| : |  |: | :|  :    :  |:|:|| ::
Db       2128 LKIKKITAIITQGCKSLSSEMYVKSYTIHYSEQGVEWKPYRLKSSMVDKIFEGNTNTKGH 2187

Qy        293 VVNSLDPPLLTRYLRIHPQSW 313
              | |  :||:::|::|: |::|
Db       2188 VKNFFNPPIISRFIRVIPKTW 2208


RESULT 14
FA5_PIG
ID   FA5_PIG         STANDARD;      PRT;  2258 AA.
AC   Q9GLP1;
DT   16-OCT-2001 (Rel. 40, Created)
```

```
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Coagulation factor V precursor (Activated protein C cofactor).
GN   Name=F5;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A., AND 3D-STRUCTURE MODELING OF F5/8 TYPE A AND C
RP   DOMAINS.
RC   TISSUE=Liver;
RX   MEDLINE=21121490; PubMed=11229814;
RA   Grimm D.R., Colter M.B., Braunschweig M., Alexander L.J., Neame P.J.,
RA   Kim H.K.W.;
RT   "Porcine factor V: cDNA cloning, gene mapping, three-dimensional
RT   protein modeling of membrane binding sites and comparative anatomy of
RT   domains.";
RL   Cell. Mol. Life Sci. 58:148-159(2001).
CC   -!- FUNCTION: Coagulation factor V is a cofactor that participates
CC       with factor Xa to activate prothrombin to thrombin.
CC   -!- SUBUNIT: Factor Va is composed of a heavy chain and a light chain,
CC       noncovalently bound. The interaction between the two chains is
CC       calcium-dependent.
CC   -!- DOMAIN: Domain B contains 41 X 9 AA tandem repeats. Domains C1 and
CC       C2 may be involved in membrane binding.
CC   -!- PTM: Thrombin activates factor V proteolytically to the active
CC       cofactor, factor Va (formation of a heavy chain at the N-terminus
CC       and a light chain at the C-terminus).
CC   -!- SIMILARITY: Belongs to the multicopper oxidase family.
CC   -!- SIMILARITY: Contains 3 F5/8 type A domains.
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF191308; AAG28381.1; -.
DR   HSSP; P12259; 1CZT.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR008979; Gal_bind_like.
DR   InterPro; IPR009271; LSPR.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   Pfam; PF06049; LSPR; 40.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
KW   Blood coagulation; Calcium; Glycoprotein; Repeat; Signal; Sulfation;
KW   Zymogen.
FT   SIGNAL        1     22       Potential.
FT   CHAIN        23   2258       Coagulation factor V.
FT   CHAIN        23    737       Coagulation factor V heavy chain (By
FT                                similarity).
FT   PROPEP      738   1611       Activation peptide (connecting region)
FT                                (By similarity).
FT   CHAIN      1612   2258       Coagulation factor V light chain (By
FT                                similarity).
FT   DOMAIN       30    329       F5/8 type A 1.
FT   DOMAIN       30    193       Plastocyanin-like 1.
FT   DOMAIN      203    329       Plastocyanin-like 2.
FT   DOMAIN      348    683       F5/8 type A 2.
FT   DOMAIN      348    525       Plastocyanin-like 3.
FT   DOMAIN      535    683       Plastocyanin-like 4.
FT   DOMAIN      691   1611       B.
FT   DOMAIN     1168   1539       41 X 9 AA approximate tandem repeats of
FT                                T-L-S-P-D-L-[GS]-[HQ]-T.
FT   REPEAT     1168   1176       1.
FT   REPEAT     1177   1185       2.
FT   REPEAT     1186   1194       3.
FT   REPEAT     1195   1203       4.
FT   REPEAT     1204   1212       5.
FT   REPEAT     1213   1221       6.
FT   REPEAT     1222   1230       7.
FT   REPEAT     1231   1239       8.
FT   REPEAT     1240   1248       9.
FT   REPEAT     1249   1257       10.
FT   REPEAT     1258   1266       11.
FT   REPEAT     1267   1275       12.
FT   REPEAT     1276   1284       13.
FT   REPEAT     1285   1293       14.
FT   REPEAT     1294   1302       15.
FT   REPEAT     1303   1311       16.
FT   REPEAT     1312   1320       17.
FT   REPEAT     1321   1329       18.
FT   REPEAT     1330   1338       19.
FT   REPEAT     1339   1347       20.
FT   REPEAT     1348   1356       21.
FT   REPEAT     1357   1365       22.
FT   REPEAT     1366   1374       23.
FT   REPEAT     1375   1383       24.
FT   REPEAT     1384   1392       25.
FT   REPEAT     1393   1401       26.
FT   REPEAT     1402   1410       27.
FT   REPEAT     1411   1419       28.
FT   REPEAT     1420   1428       29.
FT   REPEAT     1429   1437       30.
FT   REPEAT     1438   1446       31.
FT   REPEAT     1447   1455       32.
FT   REPEAT     1456   1464       33.
FT   REPEAT     1465   1473       34.
FT   REPEAT     1474   1482       35.
FT   REPEAT     1483   1491       36.
FT   REPEAT     1492   1500       37.
FT   REPEAT     1501   1509       38.
FT   REPEAT     1510   1518       39.
FT   REPEAT     1519   1527       40.
FT   REPEAT     1531   1539       41.
FT   DOMAIN     1616   1941       F5/8 type A 3.
FT   DOMAIN     1616   1785       Plastocyanin-like 5.
FT   DOMAIN     1795   1941       Plastocyanin-like 6.
FT   DOMAIN     1941   2095       F5/8 type C 1.
FT   DOMAIN     2100   2255       F5/8 type C 2.
FT   SITE        737    738       Cleavage (by thrombin) (By similarity).
FT   SITE       1029   1030       Cleavage (by thrombin) (By similarity).
FT   SITE       1611   1612       Cleavage (by thrombin) (By similarity).
FT   DISULFID    167    193       By similarity.
FT   DISULFID    248    329       By similarity.
FT   DISULFID    499    525       By similarity.
FT   DISULFID    602    683       By similarity.
FT   DISULFID   1759   1785       Probable.
FT   DISULFID   1941   2095       By similarity.
FT   DISULFID   2100   2255       By similarity.
FT   MOD_RES     692    692       Sulfotyrosine (Potential).
FT   MOD_RES     696    696       Sulfotyrosine (Potential).
FT   MOD_RES     724    724       Sulfotyrosine (Potential).
FT   MOD_RES     726    726       Sulfotyrosine (Potential).
FT   MOD_RES     745    745       Sulfotyrosine (Potential).
FT   MOD_RES    1560   1560       Sulfotyrosine (Potential).
FT   MOD_RES    1576   1576       Sulfotyrosine (Potential).
FT   MOD_RES    1581   1581       Sulfotyrosine (Potential).
FT   MOD_RES    1584   1584       Sulfotyrosine (Potential).
FT   MOD_RES    1588   1588       Sulfotyrosine (Potential).
FT   MOD_RES    1631   1631       Sulfotyrosine (Potential).
FT   CARBOHYD    153    153       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    225    225       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    239    239       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    297    297       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    382    382       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    460    460       N-linked (GlcNAc. . .) (Potential).
```

```
FT   CARBOHYD     467     467       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     553     553       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     741     741       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     752     752       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     760     760       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     776     776       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     782     782       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     899     899       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     960     960       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1048    1048       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1057    1057       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1066    1066       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1073    1073       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1089    1089       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1174    1174       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1480    1480       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1537    1537       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1597    1597       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1737    1737       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    1886    1886       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    2044    2044       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    2243    2243       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   2258 AA;  256078 MW;  9159B9E0076A2ACC CRC64;

  Query Match             44.6%;  Score 743.5;  DB 1;  Length 2258;
  Best Local Similarity   44.4%;  Pred. No. 8.4e-54;
  Matches  143;  Conservative   63;  Mismatches  105;  Indels   11;  Gaps    4;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||:  ||| | ||:   :|: |:|:::| | | || ||  :| | ||||||:  || ||
Db       1922 VGENQRAGMQTPFLIIDRECKMPMGLSTGLIADSQIKASEFWGHWQPKLARLNNGGSYNA 1981

Qy         62 WSTKEPFS-------WIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQT 114
              | | : ||       ||:||:   ::  ||:||||:    | | ::| : || | : |:
Db       1982 WIT-DKFSGESNSKPWIQVDMQREVVFTGIQTQGAKYYLKSYYTTEFNVAYSSDQRNWRI 2040

Qy        115 YRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSC 174
              ::|||  :| | || |:| |  | |:||::|||||: ||    :  ||:|| ||::| |
Db       2041 FKGNSTKNVMYFNGNSDASTITENQFDPPVVARYIRISPTESYNKPALRLELQGCEVNGC 2100

Qy        175 SMPLGMESKAISDAQITASSYFTNMFAT-WSPSKARLHLQGRSNAWRPQVNNPKEWLQVD 233
              | ||||||  | : ||||||:  : :   | | :|||: ||| |||: : ||  :|||:|
Db       2101 STPLGMESGNIKNEQITASSFKKSWWGDYWEPFRARLNAQGRVNAWQAKANNNNQWLQID 2160

Qy        234 FQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--KVFQGNQDSFT 291
              |  |:| :|||| ||| : |||| : |  |  | :|  : :   :  |:|:|| :
Db       2161 LLKIKKITAITTQGCKSLSSEMYVKRYTIQYSDRGVEWKSYREKSSMVDKIFEGNNNIKG 2220

Qy        292 PVVNSLDPPLLTRYLRIHPQSW 313
              | |  :||:::|::||  |: |
Db       2221 HVKNFFNPPIISRFIRIIPKMW 2242


RESULT 15
O88783
ID   O88783      PRELIMINARY;      PRT;  2183 AA.
AC   O88783;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Coagulation factor V.
GN   Name=F5;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98282202; PubMed=9616155;
RA   Yang T.L., Cui J., Rehumtulla A., Yang A., Moussalli M., Kaufman R.J.,
RA   Ginsburg D.;
RT   "The structure and function of murine factor V and its inactivation by
RT   protein C.";
RL   Blood 91:4593-4599(1998).
CC   -!- SIMILARITY: Contains 2 F5/8 type C domains.
DR   EMBL; U52925; AAC99553.1; -.
DR   PIR; T42764; T42764.
DR   HSSP; P12259; 1CZT.
DR   MGD; MGI:88382; F5.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001117; Cu-oxidase.
DR   InterPro; IPR008972; Cupredoxin.
DR   InterPro; IPR000421; FA58_C.
DR   InterPro; IPR009271; LSPD.
DR   Pfam; PF00754; F5_F8_type_C; 2.
DR   Pfam; PF06049; LSPR; 24.
DR   SMART; SM00231; FA58C; 2.
DR   PROSITE; PS01285; FA58C_1; 2.
DR   PROSITE; PS01286; FA58C_2; 2.
DR   PROSITE; PS50022; FA58C_3; 2.
DR   PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_2.
SQ   SEQUENCE   2183 AA;  247228 MW;  BF0A8AA723F60317 CRC64;

  Query Match             44.5%;  Score 741.5;  DB 2;  Length 2183;
  Best Local Similarity   45.2%;  Pred. No. 1.2e-53;
  Matches  145;  Conservative   60;  Mismatches  107;  Indels    9;  Gaps    4;

Qy          2 IGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASGQYGQWAPKLARLHYSGSINA 61
              :||:  ||| | ||:   :|: |:|:::| | | || ||      | |:||||: :|| ||
Db       1847 VGENQVAGMQTPFLIIDKECKMPMGLSTGVISDSQIKASEYLTYWEPRLARLNNAGSYNA 1906

Qy         62 WSTKE-----PFS-WIKVDLLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTY 115
              || ::     |   ||:||:   ::: ||:||||:    | : ::| : || |   || :
Db       1907 WSIEKTALDFPIKPWIQVDMQKEVVVTGIQTQGAKHYLKSCFTTEFQVAYSSDQTNWQIF 1966

Qy        116 RGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTLRMELMGCDLNSCS 175
              || |  ::| | || | | || |  :|||:|||||:|||    | |||:|| ||::| ||
Db       1967 RGKSGKSVMYFTGNSDGSTIKENRLDPPIVARYIRIHPTKSYNRPTLRLELQGCEVNGCS 2026

Qy        176 MPLGMESKAISDAQITASSYFTNMFAT-WSPSKARLHLQGRSNAWRPQVNNPKEWLQVDF 234
              |||:|   | | ||||||:  : :   | || |||: ||| |||: : || |:|||||
Db       2027 TPLGLEDGRIQDKQITASSFKKSWWGDYWEPSLARLNAQGRVNAWQAKANNNKQWLQVDL 2086

Qy        235 QKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--KVFQGNQDSFTP 292
              |  ||| : ||| ||| : |||| : |  |  |  |  : |   :  |:|:|| ::
Db       2087 LKIKKVTAIVTQGCKSLSSEMYVKSYSIQYSDQGVAWKPYRQKSSMVDKIFEGNSNTKGH 2146

Qy        293 VVNSLDPPLLTRYLRIHPQSW 313
              : | :||:::|::||  |::|
Db       2147 MKNFFNPPIISRFIRIIPKTW 2167


Search completed: April 20, 2005, 17:21:03
Job time : 80.25 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:08:58 ; Search time 28.3812 Seconds
(without alignments)
1240.089 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2242_2332
Perfect score: 484
Sequence: 1 VNNPKEWLQVDFQKTMKVTG.....VNSLDPPLLTRYLRIHPQSW 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	484	100.0	160	2	AAW69527	Aaw69527 C2 peptid
2	484	100.0	216	8	ADQ39370	Adq39370 Human myo
3	484	100.0	642	6	ABP60515	Abp60515 Human fac
4	484	100.0	643	6	ABP60514	Abp60514 Human fac
5	484	100.0	684	2	AAR73022	Aar73022 Human Fac
6	484	100.0	684	2	AAR74091	Aar74091 Factor-VI
7	484	100.0	770	3	ADP20767	Adp20767 Factor VI
8	484	100.0	790	5	ADE64594	Ade64594 Recombina
9	484	100.0	1014	1	AAP71139	Aap71139 Facor VII
10	484	100.0	1383	2	AAW33227	Aaw33227 Procoagul
11	484	100.0	1383	2	AAW33228	Aaw33228 Procoagul
12	484	100.0	1383	2	AAW33229	Aaw33229 Procoagul
13	484	100.0	1421	8	ADQ37613	Adq37613 Human Fac
14	484	100.0	1421	8	ADQ37598	Adq37598 Human Fac
15	484	100.0	1422	8	ADQ37591	Adq37591 Human Fac
16	484	100.0	1422	8	ADQ37606	Adq37606 Human Fac
17	484	100.0	1424	1	AAP80268	Aap80268 Modified
18	484	100.0	1424	1	AAP91169	Aap91169 Sequence
19	484	100.0	1424	4	AAB48842	Aab48842 Mutant ma
20	484	100.0	1424	5	AAO18622	Aao18622 Human mat
21	484	100.0	1425	1	AAP80267	Aap80267 Modified
22	484	100.0	1425	8	ADQ37605	Adq37605 Human Fac
23	484	100.0	1425	8	ADQ37590	Adq37590 Human Fac
24	484	100.0	1428	8	ADQ37599	Adq37599 Human Fac
25	484	100.0	1428	8	ADQ37604	Adq37604 Human Fac
26	484	100.0	1428	8	ADQ37589	Adq37589 Human Fac
27	484	100.0	1428	8	ADQ37614	Adq37614 Human Fac
28	484	100.0	1434	8	ADQ37594	Adq37594 Human Fac
29	484	100.0	1434	8	ADQ37609	Adq37609 Human Fac
30	484	100.0	1435	8	ADQ37600	Adq37600 Human Fac
31	484	100.0	1435	8	ADQ37615	Adq37615 Human Fac
32	484	100.0	1437	8	ADQ37593	Adq37593 Human Fac
33	484	100.0	1437	8	ADQ37608	Adq37608 Human Fac
34	484	100.0	1438	3	AAB01262	Aab01262 B-domain
35	484	100.0	1438	7	ADF31316	Adf31316 Variant o
36	484	100.0	1438	7	ADM75988	Adm75988 Modified
37	484	100.0	1438	7	ADM75985	Adm75985 Modified
38	484	100.0	1438	7	ADM75986	Adm75986 Modified
39	484	100.0	1438	7	ADM75989	Adm75989 Modified
40	484	100.0	1438	7	ADM75983	Adm75983 Wild-type
41	484	100.0	1438	7	ADM75984	Adm75984 Modified
42	484	100.0	1438	7	ADM75987	Adm75987 Modified
43	484	100.0	1440	2	AAR12971	Aar12971 Factor VI
44	484	100.0	1440	8	ADQ37607	Adq37607 Human Fac
45	484	100.0	1440	8	ADQ37592	Adq37592 Human Fac

ALIGNMENTS

RESULT 1
AAW69527
ID AAW69527 standard; peptide; 160 AA.
XX
AC AAW69527;
XX
DT 09-NOV-1998 (first entry)
XX
DE C2 peptide SEQ ID NO:9 from WO9829453 Example 3.
XX
KW Drug; cell membrane-directed drug; phospholipid; lipid bilayer;
KW cell cortex; blood coagulation; inflammation; immunological disorder.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9829453-A1.
XX
PD 09-JUL-1998.
XX
PF 05-JAN-1998; 98WO-JP000002.
XX
PR 27-DEC-1996; 96JP-00359053.
XX
PA (MOCH) MOCHIDA PHARM CO LTD.
XX
PI Kuriyama S, Hasegawa T;
XX
DR WPI; 1998-388051/33.
DR N-PSDB; AAV40051.
XX
PT Drugs containing peptide(s) with specific affinity to phospholipid(s) -
PT such as phosphatidyl serine, for treatment of blood coagulation,
PT inflammatory and immunological disorders.
XX
PS Example 3; Page 72-73; 117pp; Japanese.
XX
CC The present invention describes drug compositions which contain as an
CC active component a peptide which has specific affinity to particular
CC phospholipids (such as phosphatidyl ethanolamine or phosphatidyl serine),
CC especially to phospholipids which constitute a lipid bilayer of cellular
CC cortex and of which the concentration in the bilayer increases in cells
CC which are abnormal (e.g. through injury, denaturation or activation). In
CC particular, the peptide contains a sequence having phospholipid affinity
CC and a structure of formula (I): (A1)a-(A2)b-(A3)c, where (A1) is one of
CC two specific sequences (see AAW69516 and AAW69519), (A2) and (A3) are
CC TRYLRIHPQSWVHQIALR, LRYLRIHPQSWVHQIALR (see AAW69517) or MEVLGCEAQNLY
CC (see AAW69518); a = 0-5; b = 1-5, and c = 0-5. Preferred are the formulae

CC A1-A2-A3, A2-A3, A2-A2-A3, A2-A2-A2-A3 or A2-A2 (especially A2-A2-A3, A2-
CC A2-A2-A3 or A2-A2). The sequence is linked to a peptide such as a blood
CC factor, especially thrombo-modulin, urina-statin or membrane cofactor
CC protein. The drugs are used for the treatment and prevention of diseases
CC involving blood coagulation, inflammatory and immunological disorders.
CC The present sequence represents a peptide from the present invention
XX
SQ Sequence 160 AA;

Query Match 100.0%; Score 484; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 9.5e-52;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
Db 51 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 110

Qy 61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
Db 111 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 141

RESULT 2
ADQ39370
ID ADQ39370 standard; protein; 216 AA.
XX
AC ADQ39370;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human myocardial infarction-associated gene derived protein, SEQ ID 1033.
XX
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO2004058052-A2.
XX
PD 15-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US040978.
XX
PR 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Devlin JJ, Iakoubova O;
XX
DR WPI; 2004-533949/51.
DR N-PSDB; ADQ38542.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 10; SEQ ID NO 1033; 145pp; English.
XX
CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
SQ Sequence 216 AA;

Query Match 100.0%; Score 484; DB 8; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.4e-51;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
Db 107 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 166

Qy 61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
Db 167 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 197

RESULT 3
ABP60515
ID ABP60515 standard; protein; 642 AA.
XX
AC ABP60515;
XX
DT 28-MAR-2003 (first entry)
XX
DE Human factor VIII light chain A2201 deletion mutant.
XX
KW Human; factor VIII; T-cell; haemostatic; gene therapy;
KW thrombotic disorder; haemophilia; mutant; mutein.
XX
OS Homo sapiens.
XX
PN WO200298454-A2.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-BE000088.
XX
PR 31-MAY-2001; 2001WO-EP006297.
PR 14-NOV-2001; 2001GB-00027321.
XX
PA (COLL-) COLLEN RES FOUND VZW ONDERWIJSEN NAVORSI.
XX
PI Jacquemin MG, Saint-Remy JR;
XX
DR WPI; 2003-140553/13.
DR N-PSDB; ABV99882.
XX
PT New recombinant modified functional polypeptide, useful as a medicament
PT for treating or preventing hemophilia A or B, comprises multiple point-
PT mutations responsible for T-cell activation reduction or elimination.
XX
PS Claim 90; Page 78-81; 85pp; English.
XX
CC The invention relates to a novel recombinant modified functional
CC polypeptide which exerts at least one function of a mammalian protein and
CC which eliminates or reduces by at least about 80%, with respect to
CC activation by the unmodified polypeptide, the activation of at least one
CC T-cell clone derived from a mammal with antibody against the unmodified

```
CC   polypeptide. The polypeptide of the invention has haemostatic activity.
CC   The polynucleotide encoding the polypeptide may have a use in gene
CC   therapy. The recombinant modified functional polypeptide is useful as a
CC   medicament for treating or preventing a thrombotic disorder, such as
CC   haemophilia A or B. The peptide is also useful for in vitro or in vivo
CC   evaluation and/or modulation of T-cell reactivity. The recombinant point-
CC   mutated factor VIII light chain molecule is used as a template for
CC   determining whether modification of the factor VIII molecule is able to
CC   reduce the clearance, e.g. to increase the half-time, of factor VIII in
CC   plasma. The present sequence represents the human factor VIII light chain
CC   A2201 deletion mutant
XX
SQ   Sequence 642 AA;

  Query Match             100.0%;  Score 484;  DB 6;  Length 642;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-51;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 592

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db        593 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 623


RESULT 4
ABP60514
ID   ABP60514 standard; protein; 643 AA.
XX
AC   ABP60514;
XX
DT   28-MAR-2003  (first entry)
XX
DE   Human factor VIII light chain P2153Q mutant.
XX
KW   Human; factor VIII; T-cell; haemostatic; gene therapy; light chain;
KW   thrombotic disorder; haemophilia; mutant; mutein.
XX
OS   Homo sapiens.
XX
PN   WO200298454-A2.
XX
PD   12-DEC-2002.
XX
PF   31-MAY-2002; 2002WO-BE000088.
XX
PR   31-MAY-2001; 2001WO-EP006297.
PR   14-NOV-2001; 2001GB-00027321.
XX
PA   (COLL-) COLLEN RES FOUND VZW ONDERWIJSEN NAVORSI.
XX
PI   Jacquemin MG,  Saint-Remy JR;
XX
DR   WPI; 2003-140553/13.
DR   N-PSDB; ABV99881.
XX
PT   New recombinant modified functional polypeptide, useful as a medicament
PT   for treating or preventing hemophilia A or B, comprises multiple point-
PT   mutations responsible for T-cell activation reduction or elimination.
XX
PS   Claim 79; Page 73-75; 85pp; English.
XX
CC   The invention relates to a novel recombinant modified functional
CC   polypeptide which exerts at least one function of a mammalian protein and
CC   which eliminates or reduces by at least about 80%, with respect to
CC   activation by the unmodified polypeptide, the activation of at least one
CC   T-cell clone derived from a mammal with antibody against the unmodified
CC   polypeptide. The polypeptide of the invention has haemostatic activity.
CC   The polynucleotide encoding the polypeptide may have a use in gene
CC   therapy. The recombinant modified functional polypeptide is useful as a
CC   medicament for treating or preventing a thrombotic disorder, such as
CC   haemophilia A or B. The peptide is also useful for in vitro or in vivo
CC   evaluation and/or modulation of T-cell reactivity. The recombinant point-
CC   mutated factor VIII light chain molecule is used as a template for
CC   determining whether modification of the factor VIII molecule is able to
CC   reduce the clearance, e.g. to increase the half-time, of factor VIII in
CC   plasma. The present sequence represents the human factor VIII light chain
CC   P2153Q mutant
XX
SQ   Sequence 643 AA;

  Query Match             100.0%;  Score 484;  DB 6;  Length 643;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-51;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 593

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db        594 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 624


RESULT 5
AAR73022
ID   AAR73022 standard; peptide; 684 AA.
XX
AC   AAR73022;
XX
DT   25-MAR-2003  (revised)
DT   21-NOV-1995  (first entry)
XX
DE   Human Factor-VIII C-terminal fragment.
XX
KW   Factor-VIII; blood-clotting; blood; coagulant; haemophilia-A.
XX
OS   Homo sapiens.
XX
PN   WO9513301-A1.
XX
PD   18-MAY-1995.
XX
PF   10-NOV-1994;   94WO-DK000424.
XX
PR   12-NOV-1993;   93DK-00001281.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Persson E;
XX
DR   WPI; 1995-194038/25.
XX
PT   Crosslinked Factor VIII polypeptide which is stable - is prepd. using
PT   bis(sulphosuccinimydyl) suberate or disuccinimydyl suberate in the
PT   presence of polysorbate 80 to produce a coagulant with long lasting
PT   activity.
XX
PS   Disclosure; Page 24; 36pp; English.
XX
CC   This is the C-terminal fragment (corresponding to AAs 1649-2332) of human
CC   Factor-VIII which may be crosslinked resulting in increased stability and
CC   retention of high activity over extended periods of time after activation
CC   by thrombin. The polypeptide is used to prevent or treat diseases caused
CC   by the absence or deficiency of Factor-VIII in a subject such as
CC   haemophilia. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 684 AA;

  Query Match             100.0%;  Score 484;  DB 2;  Length 684;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-51;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          575 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 634

Qy           61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
                ||||||||||||||||||||||||||||||
Db          635 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 665


RESULT 6
AAR74091
ID   AAR74091 standard; protein; 684 AA.
XX
AC   AAR74091;
XX
DT   25-MAR-2003  (revised)
DT   04-NOV-1995  (first entry)
XX
DE   Factor-VIII light chain C-terminal fragment.
XX
KW   human; Factor VIII; light chain; C-terminal fragment; thrombin cleavage;
KW   blood-clotting.
XX
OS   Homo sapiens.
XX
PN   WO9513300-A1.
XX
PD   18-MAY-1995.
XX
PF   10-NOV-1994;   94WO-DK000423.
XX
PR   12-NOV-1993;   93DK-00001280.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Kjalke M,  Ezban Rasmussen M;
XX
DR   WPI; 1995-194037/25.
XX
PT   Factor VIII polypeptide(s) comprising a heavy chain shorter than native
PT   A1-A2 domain - are easier to produce recombinantly and retain coagulant
PT   activity, may be used to treat patients who have developed antibodies to
PT   C-terminal epitope(s) of Factor VIII.
XX
PS   Claim 4; Page 33-35; 51pp; English.
XX
CC   The sequence represents C-terminal residues 1649-2332 of a human Factor-
CC   VIII light chain. The sequence is produced by treating a full-length
CC   Factor-VIII polypeptide with a protease, e.g. thrombin. The fragment may
CC   be produced recombinantly in conjunction with a C-terminally truncated
CC   heavy chain to reduce production costs and improve safety, and production
CC   levels and stability are higher than for the full-length form. When the
CC   fragment is used along with a C-terminally truncated heavy chain, it may
CC   be used to treat patients who have developed antibodies against epitopes
CC   in the C-terminal part of the heavy chain. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 684 AA;

  Query Match             100.0%;  Score 484;  DB 2;  Length 684;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-51;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         575 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 634

Qy          61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
                ||||||||||||||||||||||||||||||
Db         635 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 665


RESULT 7
ADP20767
ID   ADP20767 standard; protein; 770 AA.
XX
AC   ADP20767;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Factor VIII protein.
XX
KW   A-domain; C-domain; blood coagulation; factor VIII;
KW   von Willebrand factor.
XX
OS   Unidentified.
XX
PN   KR99066381-A.
XX
PD   16-AUG-1999.
XX
PF   24-JAN-1998;   98KR-00002264.
XX
PR   24-JAN-1998;   98KR-00002264.
XX
PA   (GREC ) KOREA GREEN CROSS CORP.
XX
PI   Kim HC,  Byun TH,  Kim SM,  Kim JM,  Lee SG;
XX
DR   WPI; 2000-547435/50.
DR   N-PSDB; ADP20766.
XX
PT   Cell line expressing A-domain, C-domain of blood coagulation factor VIII
PT   and von Willebrand factor simultaneously, method - for manufacturing
PT   recombinant blood coagulation factor VIII using the same.
XX
PS   Disclosure; Fig 2; 21pp; Korean.
XX
CC   The invention relates to a novel cell line expressing A-domain, C-domain
CC   of blood coagulation factor VIII and von Willebrand factor
CC   simultaneously, and a method for manufacturing recombinant blood
CC   coagulation factor VIII using the same. The present sequence represents a
CC   factor VIII protein.
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 484;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 8e-51;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 720

Qy          61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
                ||||||||||||||||||||||||||||||
Db         721 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 751


RESULT 8
ADE64594
ID   ADE64594 standard; protein; 790 AA.
XX
AC   ADE64594;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Recombinant blood coagulation factor VIII protein, SEQ ID 4.
XX
KW   blood coagulation factor VIII; type-A haemophilia.
XX
OS   Unidentified.
XX
PN   CN1361178-A.
XX
PD   31-JUL-2002.
XX
```

```
PF   29-DEC-2000; 2000CN-00137779.
XX
PR   29-DEC-2000; 2000CN-00137779.
XX
PA   (SHAN-) SHANGHAI BIO-CHEM INST CHINESE ACAD SCI.
XX
PI   Qi Z,  Wang Q,  Chen C;
XX
DR   WPI; 2002-741852/81.
DR   N-PSDB; ADE64593.
XX
PT   New recombinant blood coagulation factor VIII and its production process
PT   and medicinal composition.
XX
PS   Claim 2; SEQ ID NO 4; 31pp; Chinese.
XX
CC   The invention relates to a novel recombinant blood coagulation factor
CC   VIII, its production process and its medicinal composite for treating
CC   type-A haemophilia. The invention further comprises a medicinal
CC   composition containing the blood coagulation factor which promotes blood
CC   coagulation to the blood plasma of type-A haemophilia patients. This
CC   sequence represents a recombinant blood coagulation factor VIII protein
CC   of the invention.
XX
SQ   Sequence 790 AA;

  Query Match            100.0%;  Score 484;  DB 5;  Length 790;
  Best Local Similarity  100.0%;  Pred. No. 8.3e-51;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        681 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 740

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db        741 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 771


RESULT 9
AAP71139
ID   AAP71139 standard; protein; 1014 AA.
XX
AC   AAP71139;
XX
DT   25-MAR-2003  (revised)
DT   14-MAY-1991  (first entry)
XX
DE   Facor VIII:c variant.
XX
KW   Factor VIII:c; variant; proteolysis; resistance;
KW   pro coagulation activity.
XX
OS   Homo sapiens.
XX
PN   WO8707144-A.
XX
PD   03-DEC-1987.
XX
PF   29-MAY-1987;   87WO-US001299.
XX
PR   29-MAY-1986;   86US-00868410.
PR   18-NOV-1986;   86US-00932767.
PR   09-DEC-1986;   86US-00939658.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Kaufman RJ,  Pittman D,  Toole JJ;
XX
DR   WPI; 1987-348539/49.
XX
PT   New deletion and replacement variants of factor 8:c - resistant to
PT   proteolysis but retaining pro-coagulant activity, and new DNA coding
PT   sequences.
XX
PS   Disclosure; Page 1; 42pp; English.
XX
CC   The full-length human factor VIII:c cDNA has been set forth in detail in
CC   WO8501961. This sequences is an example of the formula A-X-B, wherein
CC   A=Ala1-Arg372, B=Ser1690-Tyr2332 and X=0-1316 amino acids substantially
CC   duplicative of sequences of amino acids within the sequence Arg372-
CC   Ser1690 of the full-length sequence. Here X=0; producing a Arg371-Ser1690
CC   fusion protein. One or more deletions or replacements at Arg 220, 226,
CC   279, 282, 336, 359, 1719, 1721 and Lys 325, 338 and Tyr 346 will produce
CC   variants which are more resistant to specific proteolytic cleavage
CC   compared with natural factor VIII:c. Pro-coagulant activity and thrombin
CC   activatibility have been retained. See also AAP71726-29. (Updated on 25-
CC   MAR-2003 to correct PA field.)
XX
SQ   Sequence 1014 AA;

  Query Match            100.0%;  Score 484;  DB 1;  Length 1014;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-50;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        905 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 964

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db        965 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 995


RESULT 10
AAW33227
ID   AAW33227 standard; protein; 1383 AA.
XX
AC   AAW33227;
XX
DT   21-OCT-2004  (revised)
DT   30-APR-1998  (first entry)
XX
DE   Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
KW   Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW   recombinant secretion; pro-coagulant activity; resistance;
KW   activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW   von Willebrand factor binding site; binding affinity;
KW   FVIII replacement therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .346
FT                   /note= "factor VIIIA heavy chain"
FT   Domain          1. .329
FT                   /note= "A1 domain"
FT   Domain          1. .179
FT                   /note= "plastocyanin-like domain 1"
FT   Disulfide-bond  153. .179
FT                   /note= "probable"
FT   Domain          187. .329
FT                   /note= "plastocyanin-like domain 2"
FT   Cleavage-site   372. .373
FT                   /note= "by thrombin"
FT   Domain          380. .711
FT                   /note= "A2 domain"
FT   Domain          380. .554
FT                   /note= "plastocyanin-like domain 3"
FT   Disulfide-bond  528. .554
FT                   /note= "probable"
FT   Domain          564. .711
FT                   /note= "plastocyanin-like domain 4"
FT   Region          711. .746
```

```
FT                   /note= "a spacer of the sequence
FT                   SFSQNSRHPSTRQKQFNATTIPENDIEKTDPWF AHRTPMPKIQNVSSSDLLMLL
FT                   is inserted between domains A2 and A3"
FT   Misc-difference 740
FT                   /label= R740A
FT                   /note= "wild type Arg replaced with Ala"
FT   Region          741. .1383
FT                   /note= "factor VIIIA light chain"
FT   Domain          746. .1073
FT                   /note= "A3 domain"
FT   Domain          1073. .1221
FT                   /note= "C1 domain"
FT   Domain          1226. .1378
FT                   /note= "C2 domain"
XX
PN   WO9740145-A1.
XX
PD   30-OCT-1997.
XX
PF   24-APR-1997;    97WO-US006563.
XX
PR   24-APR-1996;    96US-0016117P.
PR   15-MAY-1996;    96US-0017785P.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Kaufman RJ,  Pipe SW,  Amano K;
XX
DR   WPI; 1997-535830/49.
XX
PT   Modified human pro-coagulant active factor VIII - can be administered to
PT   haemophiliacs, i.e. factor VIII replacement therapy.
XX
PS   Claim 20; Page; 57pp; English.
XX
CC   The present sequence represents a novel pro-coagulant active factor VIII
CC   (FVIII) mutant protein, comprising a deletion of the B domain and von
CC   Willebrand factor binding site, a mutation at Arg740 and an addition of
CC   an amino acid sequence spacer between the A2 and A3 domains. Factor VIII,
CC   along with calcium and phospholipid, acts as a cofactor for factor IXA,
CC   when it converts factor X to the activated form (factor XA). FVIII is the
CC   coagultion factor deficient in the X-chromosome-linked bleeding disorder
CC   haemophilia A. Several other mutant FVIII proteins have also been created
CC   (see AAW33222-29). The FVIII mutant F309S (AAW33225) is capable of
CC   recombinant secretion at higher levels than typically obtained with wild
CC   type FVIII and retains pro-coagulant activity. The FVIII mutant R336I
CC   (AAW33222) and R562K (AAW33223) are resistant to activated protein C
CC   (APC) cleavage. The present FVIII mutant can form a more stable
CC   configuration, and have an approximate 5-fold increase in specific
CC   activity compared to purified wild type FVIII, while increasing their
CC   binding affinity to von Willebrand factor improves their stability. The
CC   FVIII proteins can be administered to haemophiliacs, i.e. FVIII
CC   replacement therapy, while the nucleic acid molecule can be used for gene
CC   therapy. note: this sequence does not appear in the specification; it was
CC   created using sequences from the given references
CC
CC   Revised record issued on 21-OCT-2004 : Correction to feature table key
XX
SQ   Sequence 1383 AA;

  Query Match             100.0%;  Score 484;  DB 2;  Length 1383;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-50;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1333

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       1334 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1364
```

```
RESULT 11
AAW33228
ID   AAW33228 standard; protein; 1383 AA.
XX
AC   AAW33228;
XX
DT   21-OCT-2004  (revised)
DT   30-APR-1998  (first entry)
XX
DE   Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
KW   Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW   recombinant secretion; pro-coagulant activity; resistance;
KW   activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW   von Willebrand factor binding site; binding affinity;
KW   FVIII replacement therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .346
FT                   /note= "factor VIIIA heavy chain"
FT   Domain          1. .329
FT                   /note= "A1 domain"
FT   Domain          1. .179
FT                   /note= "plastocyanin-like domain 1"
FT   Disulfide-bond  153. .179
FT                   /note= "probable"
FT   Domain          187. .329
FT                   /note= "plastocyanin-like domain 2"
FT   Misc-difference 336
FT                   /label= R336I
FT                   /note= "wild type Arg replaced with Ile"
FT   Cleavage-site   372. .373
FT                   /note= "by thrombin"
FT   Domain          380. .711
FT                   /note= "A2 domain"
FT   Domain          380. .554
FT                   /note= "plastocyanin-like domain 3"
FT   Disulfide-bond  528. .554
FT                   /note= "probable"
FT   Misc-difference 562
FT                   /label= R562K
FT                   /note= "wild type Arg replaced with Lys"
FT   Domain          564. .711
FT                   /note= "plastocyanin-like domain 4"
FT   Region          711. .746
FT                   /note= "a spacer of the sequence
FT                   SFSQNSRHPSTRQKQFNATTIPENDIEKTDPWF AHRTPMPKIQNVSSSDLLMLL
FT                   is inserted between domains A2 and A3"
FT   Misc-difference 740
FT                   /label= R740A
FT                   /note= "wild type Arg replaced with Ala"
FT   Region          741. .1383
FT                   /note= "factor VIIIA light chain"
FT   Domain          746. .1073
FT                   /note= "A3 domain"
FT   Domain          1073. .1221
FT                   /note= "C1 domain"
FT   Domain          1226. .1378
FT                   /note= "C2 domain"
XX
PN   WO9740145-A1.
XX
PD   30-OCT-1997.
XX
PF   24-APR-1997;    97WO-US006563.
XX
PR   24-APR-1996;    96US-0016117P.
PR   15-MAY-1996;    96US-0017785P.
XX
PA   (UNMI ) UNIV MICHIGAN.
```

```
XX
PI   Kaufman RJ,  Pipe SW,  Amano K;
XX
DR   WPI; 1997-535830/49.
XX
PT   Modified human pro-coagulant active factor VIII - can be administered to
PT   haemophiliacs, i.e. factor VIII replacement therapy.
XX
PS   Claim 18; Page; 57pp; English.
XX
CC   The present sequence represents a novel pro-coagulant active factor VIII
CC   (FVIII) mutant protein, comprising a deletion of the B domain and von
CC   Willebrand factor binding site, mutations R336I, R562K and R740A and an
CC   addition of an amino acid sequence spacer between the A2 and A3 domains.
CC   Factor VIII, along with calcium and phospholipid, acts as a cofactor for
CC   factor IXA, when it converts factor X to the activated form (factor XA).
CC   FVIII is the coagultion factor deficient in the X-chromosome-linked
CC   bleeding disorder haemophilia A. Several other mutant FVIII proteins have
CC   also been created (see AAW33222-29). The FVIII mutant F309S (AAW33225) is
CC   capable of recombinant secretion at higher levels than typically obtained
CC   with wild type FVIII and retains pro-coagulant activity. The FVIII mutant
CC   R336I (AAW33222) and R562K (AAW33223) are resistant to activated protein
CC   C (APC) cleavage. The FVIII mutant comprising a deletion of the B domain
CC   and von Willebrand factor binding site, a mutation at Arg740 and an
CC   addition of an amino acid sequence spacer between the A2 and A3 domains
CC   can form a more stable configuration, and have an approximate 5-fold
CC   increase in specific activity compared to purified wild type FVIII, while
CC   increasing their binding affinity to von Willebrand factor improves their
CC   stability. The FVIII proteins can be administered to haemophiliacs, i.e.
CC   FVIII replacement therapy, while the nucleic acid molecule can be used
CC   for gene therapy. note: this sequence does not appear in the
CC   specification; it was created using sequences from the given references
CC
CC   Revised record issued on 21-OCT-2004 : Correction to feature table key
XX
SQ   Sequence 1383 AA;

  Query Match             100.0%;  Score 484;  DB 2;  Length 1383;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-50;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1333

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       1334 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1364


RESULT 12
AAW33229
ID   AAW33229 standard; protein; 1383 AA.
XX
AC   AAW33229;
XX
DT   21-OCT-2004  (revised)
DT   30-APR-1998  (first entry)
XX
DE   Procoagulant-active human factor VIII:C (FVIII) mutant protein.
XX
KW   Pro-coagulant active factor VIII; FVIII; haemophilia A;
KW   recombinant secretion; pro-coagulant activity; resistance;
KW   activated protein C cleavage; APC; B domain; A2 domain; A3 domain;
KW   von Willebrand factor binding site; binding affinity;
KW   FVIII replacement therapy.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .346
FT                   /note= "factor VIIIA heavy chain"
FT   Domain          1. .329
FT                   /note= "A1 domain"
FT   Domain          1. .179
FT                   /note= "plastocyanin-like domain 1"
FT   Disulfide-bond  153. .179
FT                   /note= "probable"
FT   Domain          187. .329
FT                   /note= "plastocyanin-like domain 2"
FT   Misc-difference 309
FT                   /label= F309S
FT                   /note= "wild type Phe replaced with Ser"
FT   Cleavage-site   372. .373
FT                   /note= "by thrombin"
FT   Domain          380. .711
FT                   /note= "A2 domain"
FT   Domain          380. .554
FT                   /note= "plastocyanin-like domain 3"
FT   Disulfide-bond  528. .554
FT                   /note= "probable"
FT   Domain          564. .711
FT                   /note= "plastocyanin-like domain 4"
FT   Region          711. .746
FT                   /note= "a spacer of the sequence
FT                   SFSQNSRHPSTRQKQFNATTIPENDIEKTDPWF AHRTPMPKIQNVSSSDLLMLL
FT                   is inserted between domains A2 and A3"
FT   Misc-difference 740
FT                   /label= R740A
FT                   /note= "wild type Arg replaced with Ala"
FT   Region          741. .1383
FT                   /note= "factor VIIIA light chain"
FT   Domain          746. .1073
FT                   /note= "A3 domain"
FT   Domain          1073. .1221
FT                   /note= "C1 domain"
FT   Domain          1226. .1378
FT                   /note= "C2 domain"
XX
PN   WO9740145-A1.
XX
PD   30-OCT-1997.
XX
PF   24-APR-1997;   97WO-US006563.
XX
PR   24-APR-1996;   96US-0016117P.
PR   15-MAY-1996;   96US-0017785P.
XX
PA   (UNMI ) UNIV MICHIGAN.
XX
PI   Kaufman RJ,  Pipe SW,  Amano K;
XX
DR   WPI; 1997-535830/49.
XX
PT   Modified human pro-coagulant active factor VIII - can be administered to
PT   haemophiliacs, i.e. factor VIII replacement therapy.
XX
PS   Claim 19; Page; 57pp; English.
XX
CC   The present sequence represents a novel pro-coagulant active factor VIII
CC   (FVIII) mutant protein, comprising a deletion of the B domain and von
CC   Willebrand factor binding site, mutations F309S, R740A and addition of an
CC   amino acid sequence spacer between the A2 and A3 domains. Factor VIII,
CC   along with calcium and phospholipid, acts as a cofactor for factor IXA,
CC   when it converts factor X to the activated form (factor XA). FVIII is the
CC   coagultion factor deficient in the X-chromosome-linked bleeding disorder
CC   haemophilia A. Several other mutant FVIII proteins have also been created
CC   (see AAW33222-29). The FVIII mutant F309S (AAW33225) is capable of
CC   recombinant secretion at higher levels than typically obtained with wild
CC   type FVIII and retains pro-coagulant activity. The FVIII mutant R336I
CC   (AAW33222) and R562K (AAW33223) are resistant to activated protein C
CC   (APC) cleavage. The present FVIII mutant can form a more stable
CC   configuration, and have an approximate 5-fold increase in specific
CC   activity compared to purified wild type FVIII, while increasing their
CC   binding affinity to von Willebrand factor improves their stability. The
```

```
CC   FVIII proteins can be administered to haemophiliacs, i.e. FVIII
CC   replacement therapy, while the nucleic acid molecule can be used for gene
CC   therapy. note: this sequence does not appear in the specification; it was
CC   created using sequences from the given references
CC
CC   Revised record issued on 21-OCT-2004 : Correction to feature table key
XX
SQ   Sequence 1383 AA;

  Query Match             100.0%;  Score 484;  DB 2;  Length 1383;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-50;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1274 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1333

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db       1334 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1364


RESULT 13
ADQ37613
ID   ADQ37613 standard; protein; 1421 AA.
XX
AC   ADQ37613;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Human Factor VIII deletion Pro739Phe mutant dB747.
XX
KW   Human; Factor VIII; clotting factor; blood; blood disorder;
KW   haemophilia A; gene therapy; mutant; mutein; B domain.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 739
FT                   /note= "Wild-type Pro substituted by Phe"
FT   Misc-difference 747. .748
FT                   /note= "Amino acids 748-1658 of the wild-type sequence
FT                   have been deleted"
XX
PN   US2004147436-A1.
XX
PD   29-JUL-2004.
XX
PF   28-JAN-2003; 2003US-00353753.
XX
PR   28-JAN-2003; 2003US-00353753.
XX
PA   (KIMH/) KIM H.
PA   (SONG/) SONG I.
PA   (CHOI/) CHOI J W.
PA   (JANG/) JANG J.
PA   (KIMY/) KIM Y.
PA   (LEEH/) LEE H S.
PA   (BANG/) BANG Y.
PA   (KIMD/) KIM D.
XX
PI   Kim H,  Song I,  Choi JW,  Jang J,  Kim Y,  Lee HS,  Bang Y,  Kim D;
XX
DR   WPI; 2004-561406/54.
XX
PT   New Factor VIII polypeptides that are more stable than full-length Factor
PT   VIII, useful for clotting blood or for preventing or treating blood
PT   disorders, such as Hemophilia A.
XX
PS   Example 4; Page; 39pp; English.
XX
CC   The invention relates to a Factor VIII polypeptide comprising an internal
CC   deletion of one or more amino acids between 1649 and 1688 fused to any
CC   amino acid sequence in B domain from about 741-782, with reference to
CC   full-length human Factor VIII appearing as ADQ37564. Also included are a
CC   pharmaceutical composition comprising the Factor VIII polypeptide (and a
CC   pharmaceutical carrier), a lyophilised composition comprising the Factor
CC   VIII polypeptide, clotting blood in a subject (comprising contacting a
CC   clotting amount of the Factor VIII polypeptide with the blood), treating
CC   Haemophilia A in a patient (comprising administering a clotting amount of
CC   the above Factor VIII polypeptide to a patient), an isolated nucleic acid
CC   encoding the Factor VIII polypeptide, an expression vector comprising the
CC   nucleic acid operably linked to a promoter, a host cell comprising the
CC   expression vector, making the Factor VIII polypeptide (comprising
CC   culturing the cell in conditions suitable for the vector to express the
CC   polypeptide, and isolating the polypeptide) and a purified antibody
CC   specific for the Factor VIII polypeptide cited above. The internal
CC   deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
CC   758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
CC   -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single
CC   chain. The proline at 739 is replaced by another amino acid. The
CC   polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
CC   encompassing fusion sites between Asn amino acid at positions 745,757 or
CC   764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
CC   reference to the full-length human Factor VIII amino acid sequence cited
CC   above. The polypeptide comprises a formula with the following linked
CC   domains: H-S-L, where the H domain represents a polypeptide sequence
CC   comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
CC   the S domain represents a polypeptide spacing linker comprising up to
CC   about 60 amino acids and the L domain represents a polypeptide sequence
CC   comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
CC   composition and methods are useful for clotting blood or for preventing
CC   or treating blood disorders, such as Haemophilia A. The present sequence
CC   is a human factor VIII B domain deletion mutant also carrying the
CC   Pro739Phe mutation. NOTE: The present sequence is not shown in the
CC   specification but was created by the indexer using information in the
CC   specification.
XX
SQ   Sequence 1421 AA;

  Query Match             100.0%;  Score 484;  DB 8;  Length 1421;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-50;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1371

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db       1372 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1402


RESULT 14
ADQ37598
ID   ADQ37598 standard; protein; 1421 AA.
XX
AC   ADQ37598;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Human Factor VIII deletion mutant dB747.
XX
KW   Human; Factor VIII; clotting factor; blood; blood disorder;
KW   haemophilia A; gene therapy; mutant; mutein; B domain.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 747. .748
FT                   /note= "Amino acids 748-1658 of the wild-type sequence
FT                   have been deleted"
XX
PN   US2004147436-A1.
```

XX
PD 29-JUL-2004.
XX
PF 28-JAN-2003; 2003US-00353753.
XX
PR 28-JAN-2003; 2003US-00353753.
XX
PA (KIMH/) KIM H.
PA (SONG/) SONG I.
PA (CHOI/) CHOI J W.
PA (JANG/) JANG J.
PA (KIMY/) KIM Y.
PA (LEEH/) LEE H S.
PA (BANG/) BANG Y.
PA (KIMD/) KIM D.
XX
PI Kim H, Song I, Choi JW, Jang J, Kim Y, Lee HS, Bang Y, Kim D;
XX
DR WPI; 2004-561406/54.
XX
PT New Factor VIII polypeptides that are more stable than full-length Factor
PT VIII, useful for clotting blood or for preventing or treating blood
PT disorders, such as Hemophilia A.
XX
PS Claim 2; Page; 39pp; English.
XX
CC The invention relates to a Factor VIII polypeptide comprising an internal
CC deletion of one or more amino acids between 1649 and 1688 fused to any
CC amino acid sequence in B domain from about 741-782, with reference to
CC full-length human Factor VIII appearing as ADQ37564. Also included are a
CC pharmaceutical composition comprising the Factor VIII polypeptide (and a
CC pharmaceutical carrier), a lyophilised composition comprising the Factor
CC VIII polypeptide, clotting blood in a subject (comprising contacting a
CC clotting amount of the Factor VIII polypeptide with the blood), treating
CC Haemophilia A in a patient (comprising administering a clotting amount of
CC the above Factor VIII polypeptide to a patient), an isolated nucleic acid
CC encoding the Factor VIII polypeptide, an expression vector comprising the
CC nucleic acid operably linked to a promoter, a host cell comprising the
CC expression vector, making the Factor VIII polypeptide (comprising
CC culturing the cell in conditions suitable for the vector to express the
CC polypeptide, and isolating the polypeptide) and a purified antibody
CC specific for the Factor VIII polypeptide cited above. The internal
CC deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
CC 758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
CC -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single
CC chain. The proline at 739 is replaced by another amino acid. The
CC polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
CC encompassing fusion sites between Asn amino acid at positions 745,757 or
CC 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
CC reference to the full-length human Factor VIII amino acid sequence cited
CC above. The polypeptide comprises a formula with the following linked
CC domains: H-S-L, where the H domain represents a polypeptide sequence
CC comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
CC the S domain represents a polypeptide spacing linker comprising up to
CC about 60 amino acids and the L domain represents a polypeptide sequence
CC comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
CC composition and methods are useful for clotting blood or for preventing
CC or treating blood disorders, such as Haemophilia A. The present sequence
CC is a human factor VIII B domain deletion mutant. NOTE: The present
CC sequence is not shown in the specification but was created by the indexer
CC using information in the specification.
XX
SQ Sequence 1421 AA;

Query Match 100.0%; Score 484; DB 8; Length 1421;
Best Local Similarity 100.0%; Pred. No. 1.8e-50;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1312 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1371

Qy 61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
||||||||||||||||||||||||||||||||
Db 1372 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1402

RESULT 15
ADQ37591
ID ADQ37591 standard; protein; 1422 AA.
XX
AC ADQ37591;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human Factor VIII Glycosylation/deletion mutant dBN(45-56).
XX
KW Human; Factor VIII; clotting factor; blood; blood disorder;
KW haemophilia A; gene therapy; mutant; mutein; N-glycosylation.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 745. .747
FT /note= "Asn is N-glycosylated"
FT Misc-difference 745. .746
FT /note= "Amino acids 746-1655 of the wild-type sequence
FT have been deleted"
XX
PN US2004147436-A1.
XX
PD 29-JUL-2004.
XX
PF 28-JAN-2003; 2003US-00353753.
XX
PR 28-JAN-2003; 2003US-00353753.
XX
PA (KIMH/) KIM H.
PA (SONG/) SONG I.
PA (CHOI/) CHOI J W.
PA (JANG/) JANG J.
PA (KIMY/) KIM Y.
PA (LEEH/) LEE H S.
PA (BANG/) BANG Y.
PA (KIMD/) KIM D.
XX
PI Kim H, Song I, Choi JW, Jang J, Kim Y, Lee HS, Bang Y, Kim D;
XX
DR WPI; 2004-561406/54.
XX
PT New Factor VIII polypeptides that are more stable than full-length Factor
PT VIII, useful for clotting blood or for preventing or treating blood
PT disorders, such as Hemophilia A.
XX
PS Claim 2; Page; 39pp; English.
XX
CC The invention relates to a Factor VIII polypeptide comprising an internal
CC deletion of one or more amino acids between 1649 and 1688 fused to any
CC amino acid sequence in B domain from about 741-782, with reference to
CC full-length human Factor VIII appearing as ADQ37564. Also included are a
CC pharmaceutical composition comprising the Factor VIII polypeptide (and a
CC pharmaceutical carrier), a lyophilised composition comprising the Factor
CC VIII polypeptide, clotting blood in a subject (comprising contacting a
CC clotting amount of the Factor VIII polypeptide with the blood), treating
CC Haemophilia A in a patient (comprising administering a clotting amount of
CC the above Factor VIII polypeptide to a patient), an isolated nucleic acid
CC encoding the Factor VIII polypeptide, an expression vector comprising the
CC nucleic acid operably linked to a promoter, a host cell comprising the
CC expression vector, making the Factor VIII polypeptide (comprising
CC culturing the cell in conditions suitable for the vector to express the
CC polypeptide, and isolating the polypeptide) and a purified antibody
CC specific for the Factor VIII polypeptide cited above. The internal
CC deletion is amino acids 746-1649, 746-1652, 746-1655, 758-1649, 758-1652,
CC 758-1655, 765-1649, 765-1652, 765-1655, 748-1658, 755-1658, 762-1658, 769
CC -1658, 776-1658, or 783-1658. The Factor VIII polypeptide is a single

CC chain. The proline at 739 is replaced by another amino acid. The
CC polypeptide comprises tripeptide sequence (Asn-X-Thr or Asn-X-Ser)
CC encompassing fusion sites between Asn amino acid at positions 745,757 or
CC 764, and Thr or Ser amino acid at positions 1651, 1654 or 1657, with
CC reference to the full-length human Factor VIII amino acid sequence cited
CC above. The polypeptide comprises a formula with the following linked
CC domains: H-S-L, where the H domain represents a polypeptide sequence
CC comprising substantially Ala-1 through Arg-740 of the human Factor VIII,
CC the S domain represents a polypeptide spacing linker comprising up to
CC about 60 amino acids and the L domain represents a polypeptide sequence
CC comprising Arg-1689 through Tyr-2332 of the human Factor VIII. The
CC composition and methods are useful for clotting blood or for preventing
CC or treating blood disorders, such as Haemophilia A. The present sequence
CC is a human factor VIII deletion mutant where the deletion creates an N-
CC glycosylation site. NOTE: The present sequence is not shown in the
CC specification but was created by the indexer using information in the
CC specification.
XX
SQ Sequence 1422 AA;

Query Match 100.0%; Score 484; DB 8; Length 1422;
Best Local Similarity 100.0%; Pred. No. 1.8e-50;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1372

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       1373 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1403
```

Search completed: April 20, 2005, 17:19:11
Job time : 28.3812 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:14:28 ; Search time 8.78465 Seconds
(without alignments)
773.288 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2242_2332
Perfect score: 484
Sequence: 1 VNNPKEWLQVDFQKTMKVTG.....VNSLDPPLLTRYLRIHPQSW 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	484	100.0	109	1	US-08-111-939-21	Sequence 21, Appl
2	484	100.0	160	2	US-08-162-402B-15	Sequence 15, Appl
3	484	100.0	160	4	US-09-331-793-9	Sequence 9, Appli
4	484	100.0	208	4	US-09-949-016-11661	Sequence 11661, A
5	484	100.0	1438	3	US-09-209-916-1	Sequence 1, Appli
6	484	100.0	1457	4	US-09-001-039B-47	Sequence 47, Appl
7	484	100.0	1471	1	US-08-683-839B-3	Sequence 3, Appli
8	484	100.0	1661	2	US-08-882-083-2	Sequence 2, Appli
9	484	100.0	1661	2	US-08-558-107-2	Sequence 2, Appli
10	484	100.0	1661	3	US-09-243-539-2	Sequence 2, Appli
11	484	100.0	2332	1	US-07-864-004B-4	Sequence 4, Appli
12	484	100.0	2332	1	US-08-251-937A-4	Sequence 4, Appli
13	484	100.0	2332	1	US-08-212-133A-2	Sequence 2, Appli
14	484	100.0	2332	1	US-08-276-594A-2	Sequence 2, Appli
15	484	100.0	2332	1	US-08-474-503-2	Sequence 2, Appli
16	484	100.0	2332	2	US-08-670-707A-2	Sequence 2, Appli
17	484	100.0	2332	3	US-09-037-601-2	Sequence 2, Appli
18	484	100.0	2332	3	US-09-324-867-3	Sequence 3, Appli
19	484	100.0	2332	3	US-09-315-179-2	Sequence 2, Appli
20	484	100.0	2332	4	US-09-523-656-2	Sequence 2, Appli
21	484	100.0	2332	4	US-09-957-641A-2	Sequence 2, Appli
22	484	100.0	2332	5	PCT-US93-03275-4	Sequence 4, Appli
23	484	100.0	2332	5	PCT-US94-13200-2	Sequence 2, Appli
24	484	100.0	2351	1	US-08-121-202-2	Sequence 2, Appli
25	484	100.0	2351	1	US-08-366-851A-2	Sequence 2, Appli
26	484	100.0	2351	4	US-10-133-907-4	Sequence 4, Appli
27	484	100.0	2351	4	US-09-001-039B-45	Sequence 45, Appl
28	484	100.0	2351	6	5171844-2	Patent No. 5171844
29	484	100.0	2351	6	5171844-2	Patent No. 5171844
30	480	99.2	218	1	US-07-607-538C-5	Sequence 5, Appli
31	480	99.2	218	2	US-08-162-402B-5	Sequence 5, Appli
32	480	99.2	218	4	US-09-364-185-5	Sequence 5, Appli
33	475	98.1	2351	6	5422260-1	Patent No. 5422260
34	475	98.1	2351	6	5422260-1	Patent No. 5422260
35	419	86.6	130	4	US-09-331-793-1	Sequence 1, Appli
36	417	86.2	2343	3	US-09-324-867-2	Sequence 2, Appli
37	406	83.9	160	4	US-09-957-641A-20	Sequence 20, Appl
38	406	83.9	2304	3	US-09-324-867-4	Sequence 4, Appli
39	406	83.9	2319	1	US-08-212-133A-8	Sequence 8, Appli
40	406	83.9	2319	1	US-08-474-503-6	Sequence 6, Appli
41	406	83.9	2319	2	US-08-670-707A-6	Sequence 6, Appli
42	406	83.9	2319	3	US-09-037-601-6	Sequence 6, Appli
43	406	83.9	2319	3	US-09-315-179-6	Sequence 6, Appli
44	406	83.9	2319	4	US-09-523-656-28	Sequence 28, Appl
45	406	83.9	2319	5	PCT-US94-13200-6	Sequence 6, Appli

ALIGNMENTS

```
RESULT 1
US-08-111-939-21
; Sequence 21, Application US/08111939
; Patent No. 5460951
;  GENERAL INFORMATION:
;    APPLICANT:  Kawai, Shinji
;    APPLICANT:  Takeshita, Sunao
;    APPLICANT:  Okazaki, Makoto
;    APPLICANT:  Amann, Egon
;    TITLE OF INVENTION:  Bone-Related Carboxypeptidase-Like
;    TITLE OF INVENTION:  Protein and Process for its Production
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/111,939
;      FILING DATE:  26-AUG-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 324033/92
;      FILING DATE:  03-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 230029/92
;      FILING DATE:  28-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Forman, David S.
;      REGISTRATION NUMBER:  33,694
;      REFERENCE/DOCKET NUMBER:  02481.1321-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4000
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  109 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-111-939-21
```

```
  Query Match             100.0%;  Score 484;  DB 1;  Length 109;
  Best Local Similarity   100.0%;  Pred. No. 5.7e-54;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 65

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db         66 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 96


RESULT 2
US-08-162-402B-15
; Sequence 15, Application US/08162402B
; Patent No. 5972337
;  GENERAL INFORMATION:
;    APPLICANT:  CERIANI, ROBERTO L.
;    APPLICANT:  PETERSON, JERRY A.
;    APPLICANT:  LAROCCA, DAVID J.
;    TITLE OF INVENTION:  46 KDALTON HUMAN MILK FAT
;    TITLE OF INVENTION:  GLOBULE (HMFG) ANTIGEN, FRAGMENTS & FUSION PROTEIN
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pretty, Schroeder & Poplawski
;      STREET:  444 South Flower St., 19th Floor
;      CITY:  Los Angeles
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/162,402B
;      FILING DATE:  03-DEC-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amzel, Viviana
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38215
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  213-622-7700
;      TELEFAX:  213-489-4210
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  160 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-162-402B-15

  Query Match             100.0%;  Score 484;  DB 2;  Length 160;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-54;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 110

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db        111 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 141


RESULT 3
US-09-331-793-9
; Sequence 9, Application US/09331793
; Patent No. 6500646
; GENERAL INFORMATION:
;  APPLICANT:  KURIYAMA, Shinichi
;  APPLICANT:  HASEGAWA, Takashi
;  TITLE OF INVENTION:  CELL MEMBRANE DIRECTED DRUGS
;  FILE REFERENCE:  1110-253P
;  CURRENT APPLICATION NUMBER: US/09/331,793
;  CURRENT FILING DATE:  1999-06-25
;  NUMBER OF SEQ ID NOS:  67
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 9
;  LENGTH: 160
;  TYPE: PRT
;  ORGANISM: Homo Sapiens
US-09-331-793-9

  Query Match             100.0%;  Score 484;  DB 4;  Length 160;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-54;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 110

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db        111 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 141


RESULT 4
US-09-949-016-11661
; Sequence 11661, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11661
;  LENGTH: 208
;  TYPE: PRT
;  ORGANISM: Human
US-09-949-016-11661

  Query Match             100.0%;  Score 484;  DB 4;  Length 208;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-53;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 158

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db        159 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 189


RESULT 5
US-09-209-916-1
; Sequence 1, Application US/09209916
```

```
; Patent No. 6358703
; GENERAL INFORMATION:
;  APPLICANT: Cho, Myung-Sam
;  APPLICANT: Chan, Sham-Yuen
;  APPLICANT: Kelsey, William
;  APPLICANT: Yee, Helena
;  TITLE OF INVENTION: Expression System for Factor VIII
;  FILE REFERENCE: MSB-7255
;  CURRENT APPLICATION NUMBER: US/09/209,916
;  CURRENT FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1438
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Derived from
;   OTHER INFORMATION: human factor VIII sequence
US-09-209-916-1

  Query Match             100.0%;  Score 484;  DB 3;  Length 1438;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1329 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1388

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db       1389 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1419


RESULT 6
US-09-001-039B-47
; Sequence 47, Application US/09001039B
; Patent No. 6818439
;  GENERAL INFORMATION:
;    APPLICANT:  Jolly, Douglas J.
;    APPLICANT:  Chang, Stephen M.W.
;    APPLICANT:  Respess, James G.
;    APPLICANT:  DePolo, Nicholas J.
;    APPLICANT:  Hsu, David Chi-Tang
;    APPLICANT:  Ibanez, Carlos E.
;    APPLICANT:  Greengard, Judith
;    APPLICANT:  Lee, Will
;    TITLE OF INVENTION:  METHODS FOR ADMINISTRATION OF
;    TITLE OF INVENTION:  RECOMBINANTGENE DELIVERY VEHICLES FOR TREATMENT
;    TITLE OF INVENTION:  OF HEMOPHILIA AND OTHER DISORDERS
;    NUMBER OF SEQUENCES:  84
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed Intellectual Property Law Group
;      STREET:  701 Fifth Avenue, Suite 6300
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/001,039B
;      FILING DATE:  13-JAN-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  1155.005 / 930049.441C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1457 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-001-039B-47

  Query Match             100.0%;  Score 484;  DB 4;  Length 1457;
  Best Local Similarity   100.0%;  Pred. No. 2e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1407

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db       1408 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1438


RESULT 7
US-08-683-839B-3
; Sequence 3, Application US/08683839B
; Patent No. 5744326
;  GENERAL INFORMATION:
;    APPLICANT:  Ill, Charles . R. et al.
;    TITLE OF INVENTION:  Use of Viral Cis-Acting Post-Transcriptional
;    TITLE OF INVENTION:  Regulatory Sequences To Increase Expression of
;    TITLE OF INVENTION:  Intronless Genes Containing Near-Consensus Splice Sites
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/683,839B
;      FILING DATE:  11-MARCH-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Remillard, Jane E.
;      REGISTRATION NUMBER:  38,872
;      REFERENCE/DOCKET NUMBER:  TTI-138
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1471 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-683-839B-3

  Query Match             100.0%;  Score 484;  DB 1;  Length 1471;
  Best Local Similarity   100.0%;  Pred. No. 2e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1362 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1421

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       1422 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1452


RESULT 8
US-08-882-083-2
; Sequence 2, Application US/08882083
; Patent No. 5869292
;  GENERAL INFORMATION:
;    APPLICANT:  VOORBERG, Johannes J.
;    TITLE OF INVENTION:  HYBRID PROTEINS WITH MODIFIED ACTIVITY
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/882,083
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/558,107
;      FILING DATE:  13-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ISACSON, John P.
;      REGISTRATION NUMBER:  33,715
;      REFERENCE/DOCKET NUMBER:  30472/212
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1661 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-882-083-2

  Query Match             100.0%;  Score 484;  DB 2;  Length 1661;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1611

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       1612 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1642


RESULT 9
US-08-558-107-2
; Sequence 2, Application US/08558107
; Patent No. 5910481
;  GENERAL INFORMATION:
;    APPLICANT:  VOORBERG, Johannes J.
;    TITLE OF INVENTION:  HYBRID PROTEINS WITH MODIFIED ACTIVITY
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/558,107
;      FILING DATE:  13-NOV-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ISACSON, John P.
;      REGISTRATION NUMBER:  33,715
;      REFERENCE/DOCKET NUMBER:  30472/212
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1661 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-558-107-2

  Query Match             100.0%;  Score 484;  DB 2;  Length 1661;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1611

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       1612 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1642


RESULT 10
US-09-243-539-2
; Sequence 2, Application US/09243539
; Patent No. 6130203
;  GENERAL INFORMATION:
;    APPLICANT:  VOORBERG, Johannes J.
;    TITLE OF INVENTION:  HYBRID PROTEINS WITH MODIFIED ACTIVITY
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/243,539
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/558,107
;      FILING DATE:  13-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ISACSON, John P.
;      REGISTRATION NUMBER:  33,715
```

```
;       REFERENCE/DOCKET NUMBER:  30472/212
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (202)672-5300
;       TELEFAX:  (202)672-5399
;       TELEX:  904136
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1661 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-09-243-539-2

  Query Match             100.0%;  Score 484;  DB 3;  Length 1661;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1552 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1611

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db       1612 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1642


RESULT 11
US-07-864-004B-4
; Sequence 4, Application US/07864004B
; Patent No. 5364771
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    APPLICANT:  Runge, Marschall S.
;    TITLE OF INVENTION:  Hybrid Human/Porcine Factor VIII
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/864,004B
;      FILING DATE:  07 APRIL 1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU106
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6508
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapien
;      TISSUE TYPE:  Liver cDNA sequence
US-07-864-004B-4

  Query Match             100.0%;  Score 484;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 12
US-08-251-937A-4
; Sequence 4, Application US/08251937A
; Patent No. 5583209
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    APPLICANT:  Runge, Marschall S.
;    TITLE OF INVENTION:  Hybrid Human/Porcine Factor VIII
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/251,937A
;      FILING DATE:  31-MAY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/864,004
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pratt, John S.
;      REGISTRATION NUMBER:  29,476
;      REFERENCE/DOCKET NUMBER:  EMU106DIV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-815-6367
;      TELEFAX:  404-815-6555
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapien
;      TISSUE TYPE:  Liver cDNA sequence
US-08-251-937A-4

  Query Match             100.0%;  Score 484;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
```

```
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 13
US-08-212-133A-2
; Sequence 2, Application US/08212133A
; Patent No. 5663060
;  GENERAL INFORMATION:
;    APPLICANT:  Lollar, John S.
;    APPLICANT:  Runge, Marschall S.
;    TITLE OF INVENTION:  Hybrid Human/Animal Factor VIII
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,133A
;      FILING DATE:  March 11, 1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/864,004
;      FILING DATE:  07-APR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  EMU/76677
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  404-572-6508
;      TELEFAX:  404-572-6555
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapien
;      TISSUE TYPE:  Liver cDNA sequence
US-08-212-133A-2

  Query Match             100.0%;  Score 484;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 14
US-08-276-594A-2
; Sequence 2, Application US/08276594A
; Patent No. 5693499
;  GENERAL INFORMATION:
;    APPLICANT:  YONEMURA, Hiroshi
;    APPLICANT:  TAJIMA, Yoshitaka
;    APPLICANT:  SUGAWARA, Keishin
;    APPLICANT:  MASUDA, Kenichi
;    TITLE OF INVENTION:  PROCESS FOR PREPARING HUMAN COAGULATION
;    TITLE OF INVENTION:  FACTOR VIII PROTEIN COMPLEX
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/276,594A
;      FILING DATE:  18-JUL-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/950,191
;      FILING DATE:  24-SEP-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 243262/1991
;      FILING DATE:  24-SEP-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  WEGNER, Harold C.
;      REGISTRATION NUMBER:  25,258
;      REFERENCE/DOCKET NUMBER:  74129/195/AOPA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2332 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-276-594A-2

  Query Match             100.0%;  Score 484;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 15
US-08-474-503-2
; Sequence 2, Application US/08474503
; Patent No. 5744446
;  GENERAL INFORMATION:
;    APPLICANT:  Emory University
;    TITLE OF INVENTION:  Hybrid Human/Animal Factor VIII
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  1100 Peachtree Street, Suite 2800
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30309
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/474,503
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Pratt, John S.
;       REGISTRATION NUMBER:  29,476
;       REFERENCE/DOCKET NUMBER:  EMU106CIP(3)
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  404-815-6500
;       TELEFAX:  404-815-6555
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2332 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  YES
;     ANTI-SENSE:  NO
;     FRAGMENT TYPE:  N-terminal
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapien
;       TISSUE TYPE:  Liver cDNA sequence
US-08-474-503-2

  Query Match             100.0%;  Score 484;  DB 1;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-52;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


Search completed: April 20, 2005, 17:22:32
Job time : 8.78465 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:21:09 ; Search time 19.5965 Seconds
(without alignments)
1543.422 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2242_2332
Perfect score: 484
Sequence: 1 VNNPKEWLQVDFQKTMKVTG.....VNSLDPPLLTRYLRIHPQSW 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	484	100.0	160	15	US-10-298-796-9	Sequence 9, Appli
2	484	100.0	216	17	US-10-741-600-1033	Sequence 1033, Ap
3	484	100.0	1438	13	US-10-006-091-1	Sequence 1, Appli
4	484	100.0	1438	13	US-10-047-257-1	Sequence 1, Appli
5	484	100.0	1438	14	US-10-225-900-1	Sequence 1, Appli
6	484	100.0	1459	15	US-10-239-498A-13	Sequence 13, Appl
7	484	100.0	1471	13	US-10-095-718-2	Sequence 2, Appli
8	484	100.0	1471	15	US-10-681-970-2	Sequence 2, Appli
9	484	100.0	2332	9	US-09-957-641-2	Sequence 2, Appli
10	484	100.0	2332	14	US-10-187-319-2	Sequence 2, Appli
11	484	100.0	2332	14	US-10-131-510A-2	Sequence 2, Appli
12	484	100.0	2332	15	US-10-445-235-2	Sequence 2, Appli
13	484	100.0	2332	15	US-10-360-101-229	Sequence 229, App
14	484	100.0	2332	15	US-10-239-498A-2	Sequence 2, Appli
15	484	100.0	2332	16	US-10-466-998A-1	Sequence 1, Appli
16	484	100.0	2332	16	US-10-721-997A-34	Sequence 34, Appl
17	484	100.0	2351	14	US-10-132-829-4	Sequence 4, Appli
18	484	100.0	2351	14	US-10-172-712-27	Sequence 27, Appl
19	484	100.0	2351	14	US-10-133-907-4	Sequence 4, Appli
20	484	100.0	2351	15	US-10-411-037-30	Sequence 30, Appl
21	484	100.0	2351	15	US-10-411-026-30	Sequence 30, Appl
22	484	100.0	2351	15	US-10-410-962-30	Sequence 30, Appl
23	484	100.0	2351	15	US-10-411-049-30	Sequence 30, Appl
24	484	100.0	2351	16	US-10-410-930-30	Sequence 30, Appl
25	484	100.0	2351	16	US-10-410-997-30	Sequence 30, Appl
26	484	100.0	2351	16	US-10-411-012-30	Sequence 30, Appl
27	484	100.0	2351	16	US-10-287-994-30	Sequence 30, Appl
28	484	100.0	2351	16	US-10-410-913-30	Sequence 30, Appl
29	484	100.0	2351	17	US-10-741-600-1034	Sequence 1034, Ap
30	484	100.0	2351	17	US-10-410-980-30	Sequence 30, Appl
31	480	99.2	218	15	US-10-038-252-5	Sequence 5, Appli
32	480	99.2	1459	15	US-10-239-498A-4	Sequence 4, Appli
33	480	99.2	1459	15	US-10-239-498A-15	Sequence 15, Appl
34	419	86.6	130	15	US-10-298-796-1	Sequence 1, Appli
35	417	86.2	1431	13	US-10-095-718-4	Sequence 4, Appli
36	417	86.2	1431	15	US-10-681-970-4	Sequence 4, Appli
37	406	83.9	2319	14	US-10-187-319-6	GENERAL INFORMA
38	406	83.9	2319	14	US-10-131-510A-6	Sequence 6, Appli
39	392	81.0	1443	14	US-10-187-319-39	Sequence 39, Appl
40	392	81.0	1443	14	US-10-131-510A-39	Sequence 39, Appl
41	392	81.0	2114	16	US-10-721-997A-33	Sequence 33, Appl
42	392	81.0	2133	14	US-10-187-319-37	Sequence 37, Appl
43	392	81.0	2133	14	US-10-131-510A-37	Sequence 37, Appl
44	222	45.9	60	15	US-10-298-796-8	Sequence 8, Appli
45	207	42.8	434	16	US-10-485-360-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-10-298-796-9
; Sequence 9, Application US/10298796
; Publication No. US20030220490A1
; GENERAL INFORMATION:
; APPLICANT: KURIYAMA, Shinichi
; APPLICANT: HASEGAWA, Takashi
; TITLE OF INVENTION: CELL MEMBRANE DIRECTED DRUGS
; FILE REFERENCE: 1110-253P
; CURRENT APPLICATION NUMBER: US/10/298,796
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US/09/331,793
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-298-796-9

Query Match 100.0%; Score 484; DB 15; Length 160;
Best Local Similarity 100.0%; Pred. No. 9.9e-50;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         51 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 110

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db        111 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 141
```

RESULT 2
US-10-741-600-1033

; Sequence 1033, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1033
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1033

Query Match 100.0%; Score 484; DB 17; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.4e-49;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
Db 107 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 166

Qy 61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
Db 167 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 197

RESULT 3
US-10-006-091-1
; Sequence 1, Application US/10006091
; Publication No. US20020102730A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.1
; CURRENT APPLICATION NUMBER: US/10/006,091
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-006-091-1

Query Match 100.0%; Score 484; DB 13; Length 1438;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
Db 1329 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1388

Qy 61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
Db 1389 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1419

RESULT 4
US-10-047-257-1
; Sequence 1, Application US/10047257
; Publication No. US20020115152A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255.2
; CURRENT APPLICATION NUMBER: US/10/047,257
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-047-257-1

Query Match 100.0%; Score 484; DB 13; Length 1438;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
Db 1329 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1388

Qy 61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
Db 1389 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1419

RESULT 5
US-10-225-900-1
; Sequence 1, Application US/10225900
; Publication No. US20030077752A1
; GENERAL INFORMATION:
; APPLICANT: Cho, Myung-Sam
; APPLICANT: Chan, Sham-Yuen
; APPLICANT: Kelsey, William
; APPLICANT: Yee, Helena
; TITLE OF INVENTION: Expression System for Factor VIII
; FILE REFERENCE: MSB-7255
; CURRENT APPLICATION NUMBER: US/10/225,900
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US/09/209,916
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1438
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Derived from
; OTHER INFORMATION: human factor VIII sequence
US-10-225-900-1

Query Match 100.0%; Score 484; DB 14; Length 1438;
Best Local Similarity 100.0%; Pred. No. 1.5e-48;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
Db 1329 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1388

Qy 61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
Db 1389 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1419

RESULT 6
US-10-239-498A-13
; Sequence 13, Application US/10239498A

```
; Publication No. US20040023333A1
; GENERAL INFORMATION:
;  APPLICANT: Hauser, Charlotte
;  APPLICANT:  Horster, Andrea
;  APPLICANT:  Schroder, Carola
;  APPLICANT:  Lehnerer, Michael
;  TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in
;  TITLE OF INVENTION:  Human Cell Lines
;  FILE REFERENCE: 80977.0001
;  CURRENT APPLICATION NUMBER: US/10/239,498A
;  CURRENT FILING DATE:  2003-07-08
;  PRIOR APPLICATION NUMBER: PCT/EP01/03220
;  PRIOR FILING DATE: 2001-03-21
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 1459
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: pTGF8-2hyg-s
US-10-239-498A-13

  Query Match            100.0%;  Score 484;  DB 15;  Length 1459;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-48;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1350 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1409

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db       1410 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1440


RESULT 7
US-10-095-718-2
; Sequence 2, Application US/10095718
; Publication No. US20020131956A1
; GENERAL INFORMATION:
;  APPLICANT: Walsh, Christopher
;  APPLICANT:  Chao, Hengjun
;  APPLICANT:  Burstein, Haim
;  APPLICANT:  Lynch, Carmel
;  APPLICANT:  Stepan, Tony
;  APPLICANT:  Munson, Keith
;  TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 35052/204375
;  CURRENT APPLICATION NUMBER: US/10/095,718
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 09/689,430
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/158,780
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1471
;   TYPE: PRT
;   ORGANISM: Homo sapiens B-domain deleted factor VIII
;   FEATURE:
;   OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-095-718-2

  Query Match            100.0%;  Score 484;  DB 13;  Length 1471;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-48;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1362 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1421

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db       1422 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1452


RESULT 8
US-10-681-970-2
; Sequence 2, Application US/10681970
; Publication No. US20040062752A1
; GENERAL INFORMATION:
;  APPLICANT: Walsh, Christopher
;  APPLICANT:  Chao, Hengjun
;  APPLICANT:  Burstein, Haim
;  APPLICANT:  Lynch, Carmel
;  APPLICANT:  Stepan, Tony
;  APPLICANT:  Munson, Keith
;  TITLE OF INVENTION: Adeno-Associated Virus Vectors Encoding Factor VIII and
;  TITLE OF INVENTION:  Methods of Using the Same
;  FILE REFERENCE: 35052/204375
;  CURRENT APPLICATION NUMBER: US/10/681,970
;  CURRENT FILING DATE:  2003-10-09
;  PRIOR APPLICATION NUMBER: US/09/689,430
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/158,780
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 1471
;   TYPE: PRT
;   ORGANISM: Homo sapiens B-domain deleted factor VIII
;   FEATURE:
;   OTHER INFORMATION: Homo sapiens BDD FVIII
US-10-681-970-2

  Query Match            100.0%;  Score 484;  DB 15;  Length 1471;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-48;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1362 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 1421

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||||||||||||||||||||||||||
Db       1422 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 1452


RESULT 9
US-09-957-641-2
; Sequence 2, Application US/09957641
; Publication No. US20020182670A1
; GENERAL INFORMATION:
;  APPLICANT: Emory University
;  TITLE OF INVENTION: MODIFIED FACTOR VIII
;  FILE REFERENCE: 75-00
;  CURRENT APPLICATION NUMBER: US/09/957,641
;  CURRENT FILING DATE:  2001-09-16
;  PRIOR APPLICATION NUMBER: US 60/234047
;  PRIOR FILING DATE: 2000-09-19
;  PRIOR APPLICATION NUMBER: US 60/236460
;  PRIOR FILING DATE: 2000-09-29
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-957-641-2

  Query Match            100.0%;  Score 484;  DB 9;  Length 2332;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-48;
```

```
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 10
US-10-187-319-2
; Sequence 2, Application US/10187319
; Publication No. US20030068785A1
;   GENERAL INFORMATION:
;        APPLICANT: Lollar, John S.
;        TITLE OF INVENTION: Hybrid Human/Animal Factor VIII
;        NUMBER OF SEQUENCES: 40
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
;             STREET: 5370 Manhattan Circle Suite 201
;             CITY: Boulder
;             STATE: Colorado
;             COUNTRY: USA
;             ZIP: 80303
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/187,319
;             FILING DATE: 27-Aug-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 09/523,656
;             FILING DATE: 2000-03-10
;             APPLICATION NUMBER: US 09/037,601
;             FILING DATE: 1998-03-10
;             APPLICATION NUMBER: WO PCT/US97/11155
;             FILING DATE: 1997-06-26
;             APPLICATION NUMBER: US 08/670,707
;             FILING DATE: 1996-06-26
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Greenlee, Lorance L.
;             REGISTRATION NUMBER: 27,894
;             REFERENCE/DOCKET NUMBER: 75-95K
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 303/499-8080
;             TELEFAX: 303/499-8089
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 2332 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: <Unknown>
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: YES
;        ANTI-SENSE: NO
;        FRAGMENT TYPE: N-terminal
;        ORIGINAL SOURCE:
;             ORGANISM: Homo sapiens
;             TISSUE TYPE: Liver
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-187-319-2

  Query Match           100.0%;  Score 484;  DB 14;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-48;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 11
US-10-131-510A-2
; Sequence 2, Application US/10131510A
; Publication No. US20030166536A1
; GENERAL INFORMATION:
;  APPLICANT: Lollar, John S
;  TITLE OF INVENTION: Modified Factor VIII
;  FILE REFERENCE: 75-95J
;  CURRENT APPLICATION NUMBER: US/10/131,510A
;  CURRENT FILING DATE:  2002-07-10
;  PRIOR APPLICATION NUMBER: U.S. 09/315,179
;  PRIOR FILING DATE: 1999-05-20
;  PRIOR APPLICATION NUMBER: U.S. 09/037,601
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: U.S. 08/670,707
;  PRIOR FILING DATE: 1996-06-26
;  PRIOR APPLICATION NUMBER: PCT/US97/11155
;  PRIOR FILING DATE: 1997-06-26
;  PRIOR APPLICATION NUMBER: PCT/US94/13200
;  PRIOR FILING DATE: 1994-11-15
;  PRIOR APPLICATION NUMBER: U.S. 08/212,133
;  PRIOR FILING DATE: 1994-03-11
;  PRIOR APPLICATION NUMBER: U.S. 07/864,004
;  PRIOR FILING DATE: 1992-04-07
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-131-510A-2

  Query Match           100.0%;  Score 484;  DB 14;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-48;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 12
US-10-445-235-2
; Sequence 2, Application US/10445235
; Publication No. US20040005670A1
; GENERAL INFORMATION:
;  APPLICANT: Katherine A. High
;  APPLICANT:  Rodney M. Camire
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;  TITLE OF INVENTION:  TREATMENT OF HEMOPHILIA A
;  FILE REFERENCE: CHOP0176
;  CURRENT APPLICATION NUMBER: US/10/445,235
;  CURRENT FILING DATE:  2003-05-22
;  PRIOR APPLICATION NUMBER: 60/382,486
;  PRIOR FILING DATE: 2002-05-22
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: homo sapiens
US-10-445-235-2
```

```
  Query Match           100.0%;  Score 484;  DB 15;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-48;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 13
US-10-360-101-229
; Sequence 229, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
;  APPLICANT: Moll, Gert  N.
;  APPLICANT:  Leenhouts, Cornelis J.
;  TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
;  FILE REFERENCE: 2183-5673
;  CURRENT APPLICATION NUMBER: US/10/360,101
;  CURRENT FILING DATE:  2003-02-07
;  PRIOR APPLICATION NUMBER: EP 02077060.8
;  PRIOR FILING DATE: 2002-05-24
;  NUMBER OF SEQ ID NOS: 309
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 229
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: sequence of factor VIII
US-10-360-101-229

  Query Match           100.0%;  Score 484;  DB 15;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-48;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 14
US-10-239-498A-2
; Sequence 2, Application US/10239498A
; Publication No. US20040023333A1
; GENERAL INFORMATION:
;  APPLICANT: Hauser, Charlotte
;  APPLICANT:  Horster, Andrea
;  APPLICANT:  Schroder, Carola
;  APPLICANT:  Lehnerer, Michael
;  TITLE OF INVENTION: Production of Recombinant Blood Clotting Factors in
;  TITLE OF INVENTION:  Human Cell Lines
;  FILE REFERENCE: 80977.0001
;  CURRENT APPLICATION NUMBER: US/10/239,498A
;  CURRENT FILING DATE:  2003-07-08
;  PRIOR APPLICATION NUMBER: PCT/EP01/03220
;  PRIOR FILING DATE: 2001-03-21
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-239-498A-2

  Query Match           100.0%;  Score 484;  DB 15;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-48;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


RESULT 15
US-10-466-998A-1
; Sequence 1, Application US/10466998A
; Publication No. US20040126856A1
; GENERAL INFORMATION:
;  APPLICANT: BAJAJ, S. Paul
;  APPLICANT:  FAY, Philip J.
;  TITLE OF INVENTION: Factor IXa: Factor VIIIa Interaction and Methods
;  TITLE OF INVENTION:  Therefor
;  FILE REFERENCE: 66153-41436
;  CURRENT APPLICATION NUMBER: US/10/466,998A
;  CURRENT FILING DATE:  2004-01-15
;  PRIOR APPLICATION NUMBER: PCT/US02/01724
;  PRIOR FILING DATE: 2002-01-23
;  PRIOR APPLICATION NUMBER: US 60/263,431
;  PRIOR FILING DATE: 2001-01-23
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 2332
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-466-998A-1

  Query Match           100.0%;  Score 484;  DB 16;  Length 2332;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-48;
  Matches   91;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2223 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2282

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2283 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2313


Search completed: April 20, 2005, 17:34:49
Job time : 19.5965 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: April 20, 2005, 17:13:08 ; Search time 6.30693 Seconds
(without alignments)
1388.270 Million cell updates/sec

Title: US-09-853-080B-49_COPY_2242_2332
Perfect score: 484
Sequence: 1 VNNPKEWLQVDFQKTMKVTG.....VNSLDPPLLTRYLRIHPQSW 91

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	484	100.0	216	2	A44258	factor VIII-associ
2	484	100.0	2351	1	EZHU	coagulation factor
3	406	83.9	2319	2	A47004	coagulation factor
4	392	81.0	2133	2	T42763	coagulation factor
5	208	43.0	427	2	JC4915	ags protein precur
6	207	42.8	463	1	A36479	milk fat globule m
7	206	42.6	2183	2	T42764	coagulation factor
8	199	41.1	2224	1	KFHU5	coagulation factor
9	198	40.9	2211	1	KFBO5	coagulation factor
10	174	36.0	409	2	T11743	pP47 protein - pig
11	174	36.0	845	2	JC5256	adipocyte transcri
12	173	35.7	719	2	S51739	transcription repr
13	164	33.9	218	2	A47285	milk fat globule p
14	159	32.9	401	2	S65138	glycoprotein antig
15	159	32.9	427	2	S74211	PAS-6/7 protein pr
16	154	31.8	927	1	JQ0948	A5 antigen precurs
17	137.5	28.4	3133	2	S52093	hemocytin - silkwo
18	127.5	26.3	737	2	T31349	hypothetical prote
19	123	25.4	1381	2	T31083	paranodin - rat
20	123	25.4	1385	2	T14158	neurexin IV - mous
21	115.5	23.9	737	2	T15615	hypothetical prote
22	112.5	23.2	913	2	A48280	receptor tyrosine
23	111.5	23.0	876	2	A49508	protein-tyrosine k
24	107.5	22.2	910	2	A53137	tyrosine kinase re
25	104.5	21.6	1283	2	T13799	neurexin IV - frui
26	103.5	21.4	819	2	I48859	tyro 10 receptor k
27	98.5	20.4	855	2	S42621	protein-tyrosine k
28	83	17.1	791	2	T16031	hypothetical prote
29	78	16.1	491	2	I40455	penicillin binding
30	73	15.1	1130	2	A48843	MHC class II trans
31	72.5	15.0	392	2	JC7633	aryl hydrocarbon n
32	71.5	14.8	157	2	H69849	hypothetical prote
33	71.5	14.8	1082	2	T42204	chromatin structur
34	70.5	14.6	589	2	T35784	probable secreted
35	69.5	14.4	441	2	C38575	gluconate 2-dehydr
36	68.5	14.2	762	2	H71916	DNA mismatch repai
37	68.5	14.2	805	2	JC7635	aryl hydrocarbon r
38	68	14.0	334	2	H75362	hypothetical prote
39	68	14.0	512	2	T47554	cytochrome P450 ho
40	67.5	13.9	221	2	C64040	hypothetical prote
41	67.5	13.9	1134	2	T20332	hypothetical prote
42	67	13.8	404	2	JN0630	microbial aspartic
43	67	13.8	853	2	AC2079	ferrichrome-iron r
44	67	13.8	2825	2	T14271	Doc4 protein, stre
45	66.5	13.7	358	2	S58376	aryl hydrocarbon r

ALIGNMENTS

RESULT 1
A44258
factor VIII-associated gene B hypothetical protein - human
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: A44258
R;Levinson, B.; Kenwrick, S.; Gamel, P.; Fisher, K.; Gitschier, J.
Genomics 14, 585-589, 1992
A;Title: Evidence for a third transcript from the human factor VIII gene.
A;Reference number: A44258; MUID:93052386; PMID:1427887
A;Accession: A44258
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-216 <LEV>
A;Cross-references: UNIPROT:Q14286; GB:M90707; NID:g182316; PIDN:AAA58466.1; PID:g182317
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F;1-53/Domain: discoidin I amino-terminal homology #status atypical <DN1>
F;57-210/Domain: discoidin I amino-terminal homology <DN2>

Query Match 100.0%; Score 484; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 2.6e-45;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 166

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db        167 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 197
```

RESULT 2
EZHU
coagulation factor VIII precursor [validated] - human
N;Alternate names: antihemophilic factor A; coagulation factor VIIIc; procoagulant compor
C;Species: Homo sapiens (man)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 09-Jul-2004
C;Accession: I54318; A00525; I58059; A23584; A26174; A42348; A43986; S63527; S66445; B42
R;Gitschier, J.; Wood, W.I.
Hum. Mol. Genet. 1, 199-200, 1992
A;Title: Sequence of the exon-containing regions of the human factor VIII gene.
A;Reference number: I54318; MUID:93265012; PMID:1303178
A;Accession: I54318
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1921,'S',1923-2351 <RES>
A;Cross-references: UNIPROT:P00451; GB:M88648; NID:g182381; PIDN:AAA52420.1; PID:g182383
R;Wood, W.I.; Capon, D.J.; Simonsen, C.C.; Eaton, D.L.; Gitschier, J.; Keyt, B.; Seeburg,
Nature 312, 330-337, 1984
A;Title: Expression of active human factor VIII from recombinant DNA clones.
A;Reference number: A00525; MUID:85061548; PMID:6438526

A;Accession: A00525
A;Molecule type: mRNA
A;Residues: 1-2351 <WOO>
A;Cross-references: EMBL:X01165; EMBL:X01166; EMBL:X01179
R;Toole, J.J.; Knopf, J.L.; Wozney, J.M.; Sultzman, L.A.; Buecker, J.L.; Pittman, D.D.; s, D.N.; Hewick, R.M.
Nature 312, 342-347, 1984
A;Title: Molecular cloning of a cDNA encoding human antihaemophilic factor.
A;Reference number: I58059; MUID:85061550; PMID:6438528
A;Accession: I58059
A;Status: nucleic acid sequence not shown; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-74,'V',76-1259,'E',1261-2351 <RE2>
A;Cross-references: GB:K01740; NID:g182802; PIDN:AAA52484.1; PID:g182803
R;Truett, M.A.; Blacher, R.; Burke, R.L.; Caput, D.; Chu, C.; Dina, D.; Hartog, K.; Kuo, .B.; Randolph, A.; Urdea, M.S.; Valenzuela, P.; Dahl, H.H.; Favalaro, J.; Hansen, J.; Nc DNA 4, 333-349, 1985
A;Title: Characterization of the polypeptide composition of human factor VIII:C and the
A;Reference number: A23584; MUID:86081164; PMID:3935400
A;Accession: A23584
A;Molecule type: mRNA
A;Residues: 1-2351 <TRU>
A;Cross-references: GB:M14113; NID:g182817; PIDN:AAA52485.1; PID:g182818
R;Eaton, D.; Rodriguez, H.; Vehar, G.A.
Biochemistry 25, 505-512, 1986
A;Title: Proteolytic processing of human factor VIII. Correlation of specific cleavages ity.
A;Reference number: A26174; MUID:86159740; PMID:3082357
A;Accession: A26174
A;Molecule type: protein
A;Residues: 20-36;392-399,'X',401-402;1668-1678;1709-1722,'D',1723-1725;1741-1755 <EAT>
R;Pittman, D.D.; Wang, J.H.; Kaufman, R.J.
Biochemistry 31, 3315-3325, 1992
A;Title: Identification and functional importance of tyrosine sulfate residues within re
A;Reference number: A42348; MUID:92207952; PMID:1554716
A;Accession: A42348
A;Molecule type: protein
A;Residues: 20-36;356-371;392-408;582-594;1668-1669,'X',1671;1672-1692;1693-1708;1709-17
A;Experimental source: recombinant material from Chinese hamster ovary cells
A;Note: sequence extracted from NCBI backbone and corrected to correspond with the publi
R;Fay, P.J.; Smudzin, T.M.
J. Biol. Chem. 264, 14005-14010, 1989
A;Title: Intersubunit fluorescence energy transfer in human factor VIII.
A;Reference number: A43986; MUID:89340500; PMID:2503509
A;Accession: A43986
A;Molecule type: protein
A;Residues: 'X',517-523;1853-1860,'X',1862-1864,'X',1866 <FAY>
R;Leyte, A.; van Schijndel, H.B.; Niehrs, C.; Huttner, W.B.; Verbeet, M.P.; Mertens, K.;
J. Biol. Chem. 266, 740-746, 1991
A;Title: Sulfation of Tyr(1680) of human blood coagulation factor VIII is essential for
A;Reference number: A56109; MUID:91093266; PMID:1898735
A;Contents: annotation; sulfation
R;Gitschier, J.; Wood, W.I.; Goralka, T.M.; Wion, K.L.; Chen, E.Y.; Eaton, D.H.; Vehar,
Nature 312, 326-330, 1984
A;Title: Characterization of the human factor VIII gene.
A;Reference number: A56196; MUID:85061547; PMID:6438525
A;Contents: annotation; introns
R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.; Hedner, U.; Ezban, M.
Protein Sci. 4, 740-746, 1995
A;Title: Locations of disulfide bonds and free cysteines in the heavy and light chains c
A;Reference number: A56216; MUID:95338127; PMID:7613471
A;Contents: annotation; disulfide bonds
A;Note: 329-Cys, 711-Cys, and 2019-Cys were shown to have free sulfhydryls
R;Kjalke, M.; Heding, A.; Talbo, G.; Persson, E.; Thomsen, J.; Ezban, M.
Eur. J. Biochem. 234, 773-779, 1995
A;Title: Amino acid residues 721-729 are required for full factor VIII activity.
A;Reference number: S63527; MUID:96163459; PMID:8575434
A;Accession: S63527
A;Molecule type: protein
A;Residues: 733-752;753-759 <KJA>
R;Lind, P.; Larsson, K.; Spira, J.; Sydow-Baeckman, M.; Almstedt, A.; Gray, E.; Sandberg
Eur. J. Biochem. 232, 19-27, 1995
A;Title: Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction
A;Reference number: S66445; MUID:96048024; PMID:7556150
A;Accession: S66445
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1668-1685 <LIN>
C;Comment: Factor VIII is activated by factor Xa and thrombin, but prolonged exposure pro
C;Genetics:
A;Gene: GDB:F8C
A;Cross-references: GDB:119124; OMIM:306700
A;Map position: Xq28-Xq28
A;Introns: 48/2; 89/1; 130/1; 201/1; 224/1; 263/1; 337/1; 424/2; 481/3; 513/1; 584/3; 63
C;Function:
A;Description: acts as a cofactor, with calcium and phospholipid, for the factor IXa prot
A;Pathway: blood coagulation
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C;Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; plas
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-2351/Product: coagulation factor VIII #status experimental <MAT>
F;20-740/Product: coagulation factor VIIIa heavy chain #status experimental <ACH>
F;20-356/Domain: A1 <DA1>
F;23-348/Domain: ferroxidase repeat homology <FO1>
F;392-759/Domain: A2 <DA2>
F;402-730/Domain: ferroxidase repeat homology <FO2>
F;760-1667/Domain: B <DB0>
F;1668-2351/Product: coagulation factor VIIIa light chain #status experimental <ACL>
F;1709-2038/Domain: A3 <DA3>
F;1716-2038/Domain: ferroxidase repeat homology <FO3>
F;2039-2191/Domain: C1 <DC1>
F;2039-2188/Domain: discoidin I amino-terminal homology <DN1>
F;2192-2351/Domain: C2 <DC2>
F;2192-2345/Domain: discoidin I amino-terminal homology <DN2>
F;60,258,601,776,803,847,919,962,982,1020,1024,1074,1085,1204,1274,1278,1301,1319,1403,1
F;172-198,267-348,547-573,649-730,1851-1877,1918-1922,2040-2188/Disulfide bonds: #status
F;355-356/Cleavage site: Arg-Met (coagulation factor Xa, protein C) #status predicted
F;365,737,738,742,1683,1699/Binding site: sulfate (Tyr) (covalent) #status experimental
F;391-392/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F;414,426/Binding site: sulfate (Tyr) (covalent) #status predicted
F;759-760/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F;1667-1668/Cleavage site: Arg-Glu (unidentified proteinase) #status experimental
F;1708-1709/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F;1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental
F;2193-2345/Disulfide bonds: #status predicted

Query Match 100.0%; Score 484; DB 1; Length 2351;
Best Local Similarity 100.0%; Pred. No. 4.7e-44;
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2242 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 2301

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |||||||||||||||||||||||||||||||
Db       2302 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 2332
```

RESULT 3
A47004
coagulation factor VIII precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: A47004
R;Elder, B.; Lakich, D.; Gitschier, J.
Genomics 16, 374-379, 1993
A;Title: Sequence of the murine factor VIII cDNA.
A;Reference number: A47004; MUID:93300511; PMID:8314577
A;Accession: A47004
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2319 <ELD>
A;Cross-references: UNIPROT:Q06194; GB:L05573; NID:g192456; PIDN:AAA37385.1; PID:g192457
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
F;1-19/Domain: signal sequence #status predicted <SIG>

F;23-349/Domain: ferroxidase repeat homology <FO1>
F;402-730/Domain: ferroxidase repeat homology <FO2>
F;1686-2006/Domain: ferroxidase repeat homology <FO3>
F;2007-2156/Domain: discoidin I amino-terminal homology <DN1>
F;2160-2313/Domain: discoidin I amino-terminal homology <DN2>

Query Match 83.9%; Score 406; DB 2; Length 2319;
Best Local Similarity 82.4%; Pred. No. 1.5e-35;
Matches 75; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

```
Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              ||:||:||||| ||||||||: ||||||| |||:||||||||||||| ||    ||||||
Db       2210 VNDPKQWLQVDLQKTMKVTGIITQGVKSLFTSMFVKEFLISSSQDGHHWTQILYNGKVKV 2269

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||| ||::||||||||||||||||| |
Db       2270 FQGNQDSSTPMMNSLDPPLLTRYLRIHPQIW 2300
```

RESULT 4
T42763
coagulation factor VIII precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42763
R;Lollar, P.
submitted to the EMBL Data Library, August 1996
A;Reference number: Z22269
A;Accession: T42763
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2133 <LOL>
A;Cross-references: UNIPROT:P12263; EMBL:U49517; NID:g1511633; PID:g1511634; PIDN:AAB067
C;Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase
C;Keywords: acute phase; blood coagulation; duplication; glycoprotein; hemophilia A; pla
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-2133/Product: coagulation factor VIII #status predicted <MAT>
F;23-349/Domain: ferroxidase repeat homology <FOX1>
F;402-730/Domain: ferroxidase repeat homology <FOX2>
F;1498-1820/Domain: ferroxidase repeat homology <FOX3>

Query Match 81.0%; Score 392; DB 2; Length 2133;
Best Local Similarity 79.1%; Pred. No. 4.5e-34;
Matches 72; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

```
Qy          1 VNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKV 60
              |:: :|||||| |||:||||:||||||||:||||||||:|||||| :|||| |:|  ||
Db       2024 VSSAEEWLQVDLQKTVKVTGITTQGVKSLLSSMYVKEFLVSSSQDGRRWTLFLQDGHTKV 2083

Qy         61 FQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||||||| |||||:||||| |||||||| ||
Db       2084 FQGNQDSSTPVVNALDPPLFTRYLRIHPTSW 2114
```

RESULT 5
JC4915
ags protein precursor - rat
N;Alternate names: O-acetyl-Gd3 ganglioside
C;Species: Rattus norvegicus (Norway rat)
C;Date: 26-Sep-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C;Accession: JC4915
R;Ogura, K.; Nara, K.; Watanabe, Y.; Kohno, K.; Tai, T.; Sanai, Y.
Biochem. Biophys. Res. Commun. 225, 932-938, 1996
A;Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.
A;Reference number: JC4915; MUID:96374422; PMID:8780713
A;Accession: JC4915
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-427 <OGU>
A;Cross-references: UNIPROT:P70490; DDBJ:D84068; NID:g1620006; PIDN:BAA12210.1; PID:g162
A;Experimental source: CST cell
C;Comment: This protein is required for the O-acetylation of disialoganglioside sialic a
C;Genetics:
A;Gene: ags
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
F;1-21/Domain: signal sequence #status predicted <SIG>
F;28-60/Domain: EGF homology <EG1>
F;68-107/Domain: EGF homology <EG2>
F;110-267/Domain: discoidin I amino-terminal homology <DN1>
F;271-427/Domain: discoidin I amino-terminal homology <DN2>

Query Match 43.0%; Score 208; DB 2; Length 427;
Best Local Similarity 43.3%; Pred. No. 7.6e-15;
Matches 39; Conservative 15; Mismatches 36; Indels 0; Gaps 0;

```
Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVF 61
              |: |||||||     ||||: ||| :      ||  : :: | || |||:: : |  |||
Db        325 NSAKEWLQVDLGTQKKVTGIITQGARDFGHIQYVASYKVAHSDDGVQWTVYEEQGTSKVF 384

Qy         62 QGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||| |: :   |  : | : ||:|: | ||
Db        385 QGNLDNNSHKKNIFEKPFMARYVRVLPLSW 414
```

RESULT 6
A36479
milk fat globule membrane protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A36479
R;Stubbs, J.D.; Lekutis, C.; Singer, K.L.; Bui, A.; Yuzuki, D.; Srinivasan, U.; Parry, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 8417-8421, 1990
A;Title: cDNA cloning of a mouse mammary epithelial cell surface protein reveals the exis
A;Reference number: A36479; MUID:91046008; PMID:2122462
A;Accession: A36479
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-463 <STU>
A;Cross-references: UNIPROT:P21956; GB:M38337; NID:g199142; PIDN:AAA39534.1; PID:g199143
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
C;Keywords: membrane protein
F;28-60/Domain: EGF homology <EG1>
F;68-107/Domain: EGF homology <EG2>
F;147-303/Domain: discoidin I amino-terminal homology <DN1>
F;307-463/Domain: discoidin I amino-terminal homology <DN2>

Query Match 42.8%; Score 207; DB 1; Length 463;
Best Local Similarity 42.2%; Pred. No. 1.1e-14;
Matches 38; Conservative 17; Mismatches 35; Indels 0; Gaps 0;

```
Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVF 61
              |: |||||||     :|||: ||| :      ||: : :: | || |||:: : |  |||
Db        361 NSAKEWLQVDLGTQRQVTGIITQGARDFGHIQYVESYKVAHSDDGVQWTVYEEQGSSKVF 420

Qy         62 QGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              ||| |: :   |  : | : ||:|: | ||
Db        421 QGNLDNNSHKKNIFEKPFMARYVRVLPVSW 450
```

RESULT 7
T42764
coagulation factor V - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42764
R;Yang, T.L.; Cui, J.; Rehumtulla, A.; Yang, A.; Moussalli, M.; Kaufman, R.J.; Ginsburg,
Blood 91, 4593-4599, 1998
A;Title: The structure and function of murine factor V and its inactivation by protein C.
A;Reference number: Z22270; MUID:98282202; PMID:9616155
A;Accession: T42764
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2183 <YAN>
A;Cross-references: UNIPROT:O88783; EMBL:U52925; NID:g3219690; PID:g3219691; PIDN:AAC995
C;Function:
A;Pathway: blood coagulation

C;Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase re
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma;
F;350-682/Domain: ferroxidase repeat homology <FOX1>
F;1541-1864/Domain: ferroxidase repeat homology <FOX2>

Query Match 42.6%; Score 206; DB 2; Length 2183;
Best Local Similarity 42.4%; Pred. No. 9e-14;
Matches 39; Conservative 19; Mismatches 32; Indels 2; Gaps 1;

```
Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--K 59
              || |:|||||  |  ||| : ||| ||| : |||| : |  |  |  |  : |   :  |
Db       2076 NNNKQWLQVDLLKIKKVTAIVTQGCKSLSSEMYVKSYSIQYSDQGVAWKPYRQKSSMVDK 2135

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :|:|| ::   : |  :||:::|::|| |::|
Db       2136 IFEGNSNTKGHMKNFFNPPIISRFIRIIPKTW 2167
```

RESULT 8
KFHU5
coagulation factor V precursor [validated] - human
N;Alternate names: coagulation labile factor; proaccelerin
C;Species: Homo sapiens (man)
C;Date: 19-May-1989 #sequence_revision 02-Jun-1995 #text_change 08-Dec-2000
C;Accession: A56172; A42344; A28028; A27498; A25897
R;Cripe, L.D.; Moore, K.D.; Kane, W.H.
Biochemistry 31, 3777-3785, 1992
A;Title: Structure of the gene for human coagulation factor V.
A;Reference number: A42344; MUID:92232668; PMID:1567832
A;Accession: A56172
A;Molecule type: DNA
A;Residues: 1-2224 <CRI>
A;Cross-references: GB:J05368
A;Accession: A42344
A;Molecule type: DNA
A;Residues: 48-58;79-89;120-130;191-201;239-249;313-323;368-378;428-437;461-471;533-542;
2070;2111-2120;2172-2181 <CR2>
R;Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Kaufm
Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987
A;Title: Complete cDNA and derived amino acid sequence of human factor V.
A;Reference number: A28028; MUID:87260886; PMID:3110773
A;Accession: A28028
A;Molecule type: mRNA
A;Residues: 1-857,'R',859-864,'R',866-924,'E',926-1763,'I',1765-2212,'T',2214-2224 <JEN>
A;Cross-references: GB:M16967
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
R;Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.
Biochemistry 26, 6508-6514, 1987
A;Title: Cloning of cDNAs coding for the heavy chain region and connecting region of hum
A;Reference number: A27498; MUID:88107560; PMID:2827731
A;Accession: A27498
A;Molecule type: mRNA
A;Residues: 1-1284,'I',1286-1600 <KAN>
A;Cross-references: GB:M17785
A;Note: parts of this sequence were determined by protein sequencing
R;Kane, W.H.; Davie, E.W.
Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986
A;Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homolog
A;Reference number: A25897; MUID:86313665; PMID:3092220
A;Accession: A25897
A;Molecule type: mRNA
A;Residues: 1188-1215,1315-2224 <KA2>
A;Cross-references: GB:M14335
A;Note: parts of this sequence were determined by protein sequencing
R;Keller, F.G.; Ortel, T.L.; Quinn-Allen, M.A.; Kane, W.H.
Biochemistry 34, 4118-4124, 1995
A;Title: Thrombin-catalyzed activation of recombinant human factor V.
A;Reference number: A56139; MUID:95210278; PMID:7696276
A;Contents: annotation; thrombin cleavage sites
C;Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
C;Genetics:
A;Gene: GDB:F5
A;Cross-references: GDB:119896; OMIM:227400
A;Map position: 1q23-1q23
A;Introns: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1; 65
C;Function:
A;Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa prote
A;Pathway: blood coagulation
C;Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase re
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma; s
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-2224/Product: coagulation factor V #status predicted <MAT>
F;29-737/Product: coagulation factor Va heavy chain #status experimental <VAH>
F;29-345/Domain: A1 <DA1>
F;33-329/Domain: ferroxidase repeat homology <FO1>
F;346-691/Domain: A2 <DA2>
F;351-684/Domain: ferroxidase repeat homology <FO2>
F;692-1573/Domain: B <DOB>
F;1183-1461/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)
F;1574-2224/Product: coagulation factor Va light chain #status experimental <VAL>
F;1574-1905/Domain: A3 <DA3>
F;1581-1905/Domain: ferroxidase repeat homology <FO3>
F;1667-1765/Region: phospholipid binding #status predicted
F;1906-2064/Domain: C1 <DC1>
F;1906-2061/Domain: discoidin I amino-terminal homology <DN1>
F;2065-2224/Domain: C2 <DC2>
F;2065-2221/Domain: discoidin I amino-terminal homology <DN2>
F;51,55,239,297,460,468,554,741,752,760,776,782,821,938,977,1074,1083,1103,1106,1479,149
F;167-193,248-329,500-526,603-684,1725-1751,1907-2061,2066-2221/Disulfide bonds: #status
F;334-335/Cleavage site: Arg-Asn (protein C) #status predicted
F;363,693,1546/Binding site: sulfate (Tyr) (covalent) #status predicted
F;376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F;382,1338/Binding site: carbohydrate (Asn) (covalent) #status absent
F;534-535/Cleavage site: Arg-Gly (protein C) #status predicted
F;737-738/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experimental
F;1046-1047/Cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experimental
F;1573-1574/Cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 41.1%; Score 199; DB 1; Length 2224;
Best Local Similarity 38.0%; Pred. No. 5.3e-13;
Matches 35; Conservative 24; Mismatches 31; Indels 2; Gaps 1;

```
Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--K 59
              || |:||::|  |  |:| : ||| ||| : |||| : |  |: | :|  :     :  |
Db       2117 NNNKQWLEIDLLKIKKITAIITQGCKSLSSEMYVKSYTIHYSEQGVEWKPYRLKSSMVDK 2176

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :|:|| ::   | |  :||:::|::|: |::|
Db       2177 IFEGNTNTKGHVKNFFNPPIISRFIRVIPKTW 2208
```

RESULT 9
KFBO5
coagulation factor V precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Mar-1993 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004
C;Accession: A42580; A36497
R;Guinto, E.R.; Esmon, C.T.; Mann, K.G.; MacGillivray, R.T.
J. Biol. Chem. 267, 2971-2978, 1992
A;Title: The complete cDNA sequence of bovine coagulation factor V.
A;Reference number: A42580; MUID:92147638; PMID:1737753
A;Accession: A42580
A;Molecule type: mRNA
A;Residues: 1-2211 <GUI>
A;Cross-references: UNIPROT:Q28107; GB:M81440; NID:g163037; PIDN:AAA30512.1; PID:g163038
A;Note: sequence extracted from NCBI backbone (NCBIN:80774, NCBIP:80776)
R;Kalafatis, M.; Jenny, R.J.; Mann, K.G.
J. Biol. Chem. 265, 21580-21589, 1990
A;Title: Identification and characterization of a phospholipid-binding site of bovine fac
A;Reference number: A36497; MUID:91072354; PMID:2254316
A;Accession: A36497
A;Molecule type: protein
A;Residues: 1566-1570,'X',1572-1581,'X',1583-1584;1673-1676,'X',1678-1679,'X',1681,'X',16
R;Xue, J.; Kalafatis, M.; Silveira, J.R.; Kung, C.; Mann, K.G.
Biochemistry 33, 13109-13116, 1994
A;Title: Determination of the disulfide bridges in factor Va heavy chain.

```
A;Reference number: A55979; MUID:95034740; PMID:7947716
A;Contents: annotation
A;Note: 566-Cys and 617-Cys were shown to have free sulfhydryls
C;Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.
C;Function:
A;Description: acts as a cofactor, with calcium and phospholipid, for the factor Xa prot
A;Pathway: blood coagulation
C;Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase re
C;Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasma;
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-2211/Product: coagulation factor V #status predicted <MAT>
F;29-741/Product: coagulation factor Va heavy chain #status predicted <VAH>
F;29-345/Domain: A1 <DA1>
F;33-329/Domain: ferroxidase repeat homology <FO1>
F;346-695/Domain: A2 <DA2>
F;351-688/Domain: ferroxidase repeat homology <FO2>
F;696-1564/Domain: B <DOB>
F;1175-1437/Region: 9-residue repeats (Q-X-T/N-L-S-P-D-L-S)
F;1565-2211/Product: coagulation factor Va light chain #status predicted <VAL>
F;1565-1892/Domain: A3 <DA3>
F;1572-1892/Domain: ferroxidase repeat homology <FO3>
F;1654-1752/Region: phospholipid binding #status predicted
F;1893-2051/Domain: C1 <DC1>
F;1893-2048/Domain: discoidin I amino-terminal homology <DN1>
F;2052-2211/Domain: C2 <DC2>
F;2052-2208/Domain: discoidin I amino-terminal homology <DN2>
F;167-193,248-329,499-525/Disulfide bonds: #status experimental
F;225,239,297,382,460,553,587,745,756,774,780,902,952,964,1044,1053,1062,1071,1078,1094,
F;334-335/Cleavage site: Arg-Asn (protein C) #status predicted
F;363,697,1537/Binding site: sulfate (Tyr) (covalent) #status predicted
F;376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted
F;533-534/Cleavage site: Arg-Gly (protein C) #status predicted
F;607-688,1712-1738,1894-2048,2053-2208/Disulfide bonds: #status predicted
F;741-742/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predicted
F;1034-1035/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predicted
F;1564-1565/Cleavage site: Arg-Ser (thrombin) #status experimental

  Query Match             40.9%;  Score 198;  DB 1;  Length 2211;
  Best Local Similarity   39.1%;  Pred. No. 6.8e-13;
  Matches   36;  Conservative   20;  Mismatches   34;  Indels    2;  Gaps    1;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--K 59
              ||  :|||:|  |  |:| : ||| ||| : |||| : |  |  |  |  : :   :  |
Db       2104 NNNNQWLQIDLLKIKKITAIVTQGCKSLSSEMYVKSYTIHYSDQGTDWKPYREKSSMVDK 2163

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :|:|| :    | |  :||:::|::|| |::|
Db       2164 IFEGNNNVRGHVKNFFNPPIISRFIRIIPKTW 2195


RESULT 10
T11743
pP47 protein - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11743
R;Ensslin, M.; Vogel, T.; Calvete, J.J.; Thole, H.H.; Schmidtke, J.; Matsuda, T.; Toepfe
Biol. Reprod. 58, 1057-1064, 1998
A;Title: Molecular cloning and characterization of P47, a novel boar sperm-associated zo
A;Reference number: Z17325; MUID:98206817; PMID:9546740
A;Accession: T11743
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-409 <ENS>
A;Cross-references: UNIPROT:P79385; EMBL:Y11683; NID:g2652927; PIDN:CAA72379.1; PID:g265
A;Experimental source: testis
C;Function:
A;Description: may be involved in membrane remodeling and/or function as a zona pellucid
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
F;6-40/Domain: EGF homology <EGF>

  Query Match             36.0%;  Score 174;  DB 2;  Length 409;
  Best Local Similarity   37.0%;  Pred. No. 3.7e-11;
  Matches   34;  Conservative   17;  Mismatches   39;  Indels    2;  Gaps    1;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKV--K 59
              |:  ||||:|      :|||: ||| :       ||  : :: | ||  ||  :   | :  |
Db        305 NSASEWLQIDLGSQRRVTGIITQGARDFGHIQYVAAYKVAYSDDGVSWTEYRDQGALEGK 364

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :| || |: :   |  : | |||::|| | :|
Db        365 IFPGNLDNNSHKKNMFETPFLTRFVRILPVAW 396


RESULT 11
JC5256
adipocyte transcription factor, AEBP1 - human
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: JC5256
R;Ohno, I.; Hashimoto, J.; Shimizu, K.; Takaoka, K.; Ochi, T.; Matsubara, K.; Okubo, K.
Biochem. Biophys. Res. Commun. 228, 411-414, 1996
A;Title: A cDNA cloning of human AEBP1 from primary cultured osteoblasts and its expressi
A;Reference number: JC5256; MUID:97079196; PMID:8920928
A;Accession: JC5256
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-845 <OHN>
A;Cross-references: UNIPROT:Q14113; DDBJ:D86479; NID:g1468942; PIDN:BAA13094.1; PID:g146

  Query Match             36.0%;  Score 174;  DB 2;  Length 845;
  Best Local Similarity   34.9%;  Pred. No. 8.8e-11;
  Matches   30;  Conservative   20;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          6 EWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQ 65
              :|::|| ::| : ||| |||  | :   :|  | :  | |   | ::    :   | ||
Db        129 QWIEVDTRRTTRFTGVITQGRDSSIHDDFVTTFFVGFSNDSQTWVMYTNGYEEMTFHGNV 188

Qy         66 DSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |  |||:: |  |:: |::||:| :|
Db        189 DKDTPVLSELPEPVVARFIRIYPLTW 214


RESULT 12
S51739
transcription repressor AEBP1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S60227; S51739
R;He, G.P.; Muise, A.; Li, A.W.; Ro, H.S.
Nature 378, 92-96, 1995
A;Title: A eukaryotic transcriptional repressor with carboxypeptidase activity.
A;Reference number: S60227; MUID:96061010; PMID:7477299
A;Accession: S60227
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-719 <HE2>
A;Cross-references: UNIPROT:Q61281; EMBL:X80478; NID:g607131; PIDN:CAA56648.1; PID:g60713

  Query Match             35.7%;  Score 173;  DB 2;  Length 719;
  Best Local Similarity   34.9%;  Pred. No. 9.4e-11;
  Matches   30;  Conservative   20;  Mismatches   36;  Indels    0;  Gaps    0;

Qy          6 EWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQ 65
              :|::|| ::| : ||| |||  | :   :|  | :  | |   | ::    :   | ||
Db         24 QWIEVDTRRTTRFTGVITQGRDSSIHDDFVTTFFVGFSNDSQTWVMYTNGYEEMTFYGNV 83

Qy         66 DSFTPVVNSLDPPLLTRYLRIHPQSW 91
              |  |||:: |  |:: |::||:| :|
Db         84 DKDTPVLSELPEPVVARFIRIYPLTW 109


RESULT 13
A47285
milk fat globule protein - human (fragment)
C;Species: Homo sapiens (man)
```

```
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A47285
R;Larocca, D.; Peterson, J.A.; Urrea, R.; Kuniyoshi, J.; Bistrain, A.M.; Ceriani, R.L.
Cancer Res. 51, 4994-4998, 1991
A;Title: A M-r 46,000 human milk fat globule protein that is highly expressed in human b
A;Reference number: A47285; MUID:91371351; PMID:1909932
A;Accession: A47285
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-218 <LAR>
A;Cross-references: UNIPROT:Q08431; GB:S56151; NID:g235396; PIDN:AAB19771.1; PID:g235397
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
F;1-56/Domain: discoidin I amino-terminal homology (fragment) <DN1>
F;60-218/Domain: discoidin I amino-terminal homology <DN2>

  Query Match             33.9%;  Score 164;  DB 2;  Length 218;
  Best Local Similarity   35.2%;  Pred. No. 2.1e-10;
  Matches   31;  Conservative   20;  Mismatches   35;  Indels    2;  Gaps    1;

Qy          6 EWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFF--QNGKVKVFQG 63
              :|||||    : :|||: ||| ::  :  :|  : :: | |    || :    : |  |:| |
Db        118 QWLQVDLGSSKEVTGIITQGARNFGSVQFVASYKVAYSNDSANWTEYQDPRTGSSKIFPG 177

Qy         64 NQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              | |: :   |  : |:| ||:|| | :|
Db        178 NWDNHSHKKNLFETPILARYVRILPVAW 205


RESULT 14
S65138
glycoprotein antigen MGP57/53, mammary gland - bovine (fragment)
N;Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFG-E8 h
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-Aug-1998
C;Accession: S65138; G48394
R;Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 385-391, 1995
A;Title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal an
A;Reference number: S65138; MUID:96125736; PMID:8541316
A;Accession: S65138
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-401 <AOK>
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A;Reference number: A48394; MUID:93250576; PMID:8485470
A;Accession: G48394
A;Status: preliminary
A;Molecule type: protein
A;Residues: 207-220 <MAT>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131457)
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
C;Keywords: glycoprotein
F;1-32/Domain: EGF homology (fragment) <EG1>
F;40-79/Domain: EGF homology <EG2>
F;82-239/Domain: discoidin I amino-terminal homology <DN1>
F;243-401/Domain: discoidin I amino-terminal homology <DN2>

  Query Match             32.9%;  Score 159;  DB 2;  Length 401;
  Best Local Similarity   33.7%;  Pred. No. 1.6e-09;
  Matches   31;  Conservative   17;  Mismatches   42;  Indels    2;  Gaps    1;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNG--KVK 59
              |:  ||||:|     :|||: ||| :       ||  : ::   ||  || :    |  : |
Db        297 NSASEWLQIDLGSQKRVTGIITQGARDFGHIQYVAAYRVAYGDDGVTWTEYKDPGASESK 356

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :| || |: :    |  : |    |::|| | :|
Db        357 IFPGNMDNNSHKKNIFETPFQARFVRIQPVAW 388


RESULT 15
S74211
PAS-6/7 protein precursor - bovine
N;Alternate names: glycoprotein component 16/major fat-globule membrane protein/MFG-E8 hc
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C;Accession: S74211; S78114; S24181; S65138; G48394
R;Hvarregaard, J.; Andersen, M.H.; Berglund, L.; Rasmussen, J.T.; Petersen, T.E.
Eur. J. Biochem. 240, 628-636, 1996
A;Title: Characterization of glycoprotein PAS-6/7 from membranes of bovine milk fat globu
A;Reference number: S74211; MUID:97008954; PMID:8856064
A;Accession: S74211
A;Molecule type: mRNA
A;Residues: 1-427 <HVA>
A;Cross-references: UNIPROT:Q95114; EMBL:X91895; NID:g1632778; PIDN:CAA62997.1; PID:g163
A;Accession: S78114
A;Molecule type: protein
A;Residues: 19-85;96-110;140-165;174-216;221-232;248-277;285-293;309-337;339-420;425-427
R;Kim, D.H.; Kanno, C.; Mizokami, Y.
Biochim. Biophys. Acta 1122, 203-211, 1992
A;Title: Purification and characterization of major glycoproteins, PAS-6 and PAS-7, from
A;Reference number: S23926; MUID:92353107; PMID:1643094
A;Accession: S24181
A;Molecule type: protein
A;Residues: 383-394 <KIM>
R;Aoki, N.; Kishi, M.; Taniguchi, Y.; Adachi, T.; Nakamura, R.; Matsuda, T.
Biochim. Biophys. Acta 1245, 385-391, 1995
A;Title: Molecular cloning of glycoprotein antigens MGP57/53 recognized by monoclonal ant
A;Reference number: S65138; MUID:96125736; PMID:8541316
A;Accession: S65138
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 27-427 <AOK>
R;Mather, I.H.; Banghart, L.R.; Lane, W.S.
Biochem. Mol. Biol. Int. 29, 545-554, 1993
A;Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig
II-like sequences.
A;Reference number: A48394; MUID:93250576; PMID:8485470
A;Accession: G48394
A;Status: preliminary
A;Molecule type: protein
A;Residues: 233-246 <MAT>
A;Experimental source: milk
A;Note: sequence extracted from NCBI backbone (NCBIP:131457)
C;Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF homolo
C;Keywords: blocked amino end; disulfide bond; glycoprotein; milk
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-427/Product: PAS-6/7 protein #status experimental <MAT>
F;24-58/Domain: EGF homology <EG1>
F;66-105/Domain: EGF homology <EG2>
F;108-265/Domain: discoidin I amino-terminal homology <DN1>
F;269-427/Domain: discoidin I amino-terminal homology <DN2>
F;24-35,29-47,49-58,66-77,71-94,96-105/Disulfide bonds: #status predicted
F;27/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;34/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;59,227/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;109-265,252-256,270-427/Disulfide bonds: #status experimental

  Query Match             32.9%;  Score 159;  DB 2;  Length 427;
  Best Local Similarity   33.7%;  Pred. No. 1.7e-09;
  Matches   31;  Conservative   17;  Mismatches   42;  Indels    2;  Gaps    1;

Qy          2 NNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNG--KVK 59
              |:  ||||:|     :|||: ||| :       ||  : ::   ||  || :    |  : |
Db        323 NSASEWLQIDLGSQKRVTGIITQGARDFGHIQYVAAYRVAYGDDGVTWTEYKDPGASESK 382

Qy         60 VFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSW 91
              :| || |: :    |  : |    |::|| | :|
Db        383 IFPGNMDNNSHKKNIFETPFQARFVRIQPVAW 414


Search completed: April 20, 2005, 17:21:43
```

Job time : 7.30693 secs